ABSTRACT

Title of Dissertation: ROBUSTNESS ANALYSIS OF SENSOR COVERAGE AND LOCATION FOR REAL-TIME TRAFFIC ESTIMATION AND PREDICTION IN LARGE-SCALE NETWORKS

Xiang Fei, Ph.D., 2007

Dissertation Directed By: Professor Hani S. Mahmassani, Department of Civil and Environmental Engineering

The growing need of agencies to obtain real-time information on the traffic state of key facilities in the systems they manage is driving interest in cost-effective deployment of sensor technologies across the networks they manage. This has led to greater interest in the sensor location problem. Finding a set of optimal sensor locations is a network design problem. This dissertation addresses a series of critical and challenging issues in the robustness analysis of sensor coverage and location under different traffic conditions, in the context of real-time traffic estimation and prediction in a large scale traffic network.

The research presented in this dissertation represents an important step towards optimization of sensor locations based on dynamic traffic assignment methodology. It proposes an effective methodology to find optimal sensor coverage and locations, for a specified number of sensors, through an iterative mathematical bi-level optimization framework, The proposed methods help transportation planners locate a minimal number of sensors to completely cover all or a subset of OD pairs in

a network without budgetary constraints, or optimally locate a limited number of sensors by considering link information gains (weights of each link brought to correct *a-priori* origin-destination flows) and flow coverage with budgetary constraints.

Network uncertainties associated with the sensor location problem are considered in the mathematical formulation. The model is formulated as a two stage stochastic model. The first stage decision denotes a strategic sensor location plan before observations of any randomness events, while the recourse function associated with the second stage denotes the expected cost of taking corrective actions to the first stage solution after the occurrence of the random events.

Recognizing the location problem as a $\mathcal{NP}$-hard problem, a hybrid Greedy Randomized Adaptive Search Procedure (GRASP) is employed to circumvent the difficulties of exhaustively exploring the feasible solutions and find a near-optimal solution for this problem. The proposed solution procedure is operated in two stages. In stage one, a restricted candidate list (RCL) is generated from choosing a set of top candidate locations sorted by the link flows. A predetermined number of links is randomly selected from the RCL according to link independent rule. In stage two, the selected candidate locations generated from stage one are evaluated in terms of the magnitude of flow variation reduction and coverage of the origin-destination flows using the archived historical and simulated traffic data. The proposed approaches are tested on several actual networks and the results are analyzed.

ROBUSTNESS ANALYSIS OF SENSOR COVERAGE AND LOCATION FOR REAL-TIME TRAFFIC ESTIMATION AND PREDICTION IN LARGE-SCALE NETWORKS.

By

Xiang Fei

Dissertation 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
Doctor of Philosophy
2007

Advisory Committee:
Professor Hani S. Mahmassani, Chair
Professor Eyad H. Abed
Professor Paul M. Schonfeld
Associate Professor Richard J. La
Assistant Professor Cinzia Cirillo

Acknowledgements

I am deeply indebted to my supervisor professor Hani S.Mahmassani for his enthusiastic encouragement, continued patience and generous support through my PhD study at University of Maryland. This work could not have been completed without his guidance, broad and deep knowledge and patience. His passionate attitude to the scientific research and outlook on life has strong influenced me on my future career path. I consider the four years working with him the most productive and exciting years of my professional life. Because of him, I have learned to read better, write better and even think better. And one of my regrets about leaving college park is the loss of such a director influence on my research and work. Professor Mahmassani was also an enormous help during my job search, serving as my motivator and chief reference to outside parties. I think that my success in being granted interviews was due in large part to his enthusiasm.

I am also thankful to Dr. Eyad H. Abed, Dr. Paul M. Schonfeld, Dr. Cinzia Cirillo and Dr. Richard J. La, members of my dissertation committee, for their participation and helps. Their valuable comments and suggestions to my initial work directly contribute to the completion of my dissertation.

The experience during my study in College park, MD would have been incomplete if it was not for the interaction I had with my fellow students, friends and colleagues. Special thanks firstly go to Dr. Xuesong Zhou for his advices during my first several years' doctoral studies. I am grateful to Dr. Xiao Qin, Dr. Chung-Cheng Jason Lu, Stacy Eisenman and Hayssam Sbayti for their valuable inputs and critical comments on the DTA system. My appreciation also goes to the other DTA research team members, Sevgi Erdogan, Jing Dong, Kuilin Zhang, and Samer Hamdar for their countless helps.

I would also like to thank Dr. Li Zhang, Dr. Wei-hua Lin, Dr. Yi-Chang Chiu, Mr. Michael Pack, Mr. Tom Jacobs, Mr. Phil Tarnoff, Dr. David Lovell and Dr. Gang-Len Chang, for their advices and assistance in the course of my research and studies.

I express my indebtedness to my grandparents, parents and sister for their love, encouragement and inspiration, and for standing by me all through my life. Lastly, but definitely not the least, I would like to extend my heartful gratitude to my beloved wife, Ke Liu, for her sincere love, compassion, encouragement, and strongly standing up for me all the way. I cannot go this far without her encouragement and unconditional support.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

1.1 Research motivation and objectives

1.1.1 Research motivation

Transportation system congestion is one of the top concerns affecting economic prosperity and people's way of life. Whatever forms it may take, such as vehicles stalled in road traffic networks, cargo stuck at overwhelmed seaports, or airplanes circling over crowded airports, congestion costs America an estimated $200 billion a year (Peters 2007). Traffic congestion leads to side effects, such as drivers' additional travel time cost on the road, extra consumption of fuel, environmental pollution, incidents, etc. As estimated by Schrank *et al* (2005), in 2003 congestion costs (based on wasted time and fuel) about $63.1 billion in the 85 urban areas, compared to $61.5 billion in 2002. The cost ranged from $1,038 per traveler in very large urban areas to $222 per traveler in smaller areas. As a rapidly developing technology, sensor networks can be part of an effective strategy to improve the overall performance of general traffic networks, contributing to the reduction of congestion and its onerous by-products. Through telecommunication and information technologies, sensor networks could form an important component of advanced traveler information systems (ATIS) that deliver traveler information to the traffic management center (TMC), and provide transportation system users with greater transportation options and travel efficiency. Improvements in sensor technology and communication systems allow transportation agencies to more closely monitor the condition of the surface transportation system and predict traffic conditions to enable proactive traffic management.

Notwithstanding continuing advances in surveillance and communication technologies, the ability to observe flow patterns and performance characteristics of dynamic transportation systems remains an important challenge for transportation agencies. As these technologies continue to become more reliable and cost effective, demand for travel information is also growing, as is the potential and ability to use sensor and probe information in sophisticated decision support systems for traffic systems management. While probe data based on cellular-assisted GPS and other cellular phone technologies hold the promise of near-ubiquitous information coverage in a network, measurements on system state at given locations using fixed sensors remain the backbone of most traffic management centers for traffic management and control purposes.

In order to improve the efficiency of data collection in transportation networks, it is critical to understand how sensor placement affects the network observability. Furthermore, a new generation of real-time network traffic estimation and prediction systems is designed to interact with real-time sensor data to support system management decisions through estimation, prediction and control generation cycles (Mahmassani et al., 2005). For example, real-time DTA systems such as DYNASMART-X and DYNAMIT use sensor measurements on a subset of the network links as basis for estimation and prediction of traffic conditions on a quasi-continuous basis. In particular, the sensor measurements are combined with current observation values and historical information to estimate prevailing origin-destination

(O-D) patterns and predict their near-term evolution, in addition to predicting the network traffic patterns associated with these O-D demands.

An OD trip table is an important input to a traffic assignment model as well as an ITS system. However, the OD demand is typically difficult to obtain due to the formulation of the demand estimation and prediction model, such as model order and model parameters, and the uncertainty associated with the demand estimation and prediction process. Substantial research has been conducted on developing demand estimation methods. Generally, demand estimation can be categorized into two classes, static and dynamic estimation. The conventional methods for collecting OD trip demand matrix information include the lights-on survey method, license plate matching method, postcard questionnaire method, and roadside destination interview method, all of which are costly, labor intensive, time consuming and disruptive. The problem becomes more acute in regions undergoing rapid development. In an attempt to circumvent these issues, many studies have been conducted on methods for analyzing collected link traffic data to estimate and predict OD demand matrices. Traffic counts are inherently attractive as data source for OD trip estimation since they are non-disruptive to travelers, generally available, and relatively inexpensive to collect. The information contained in time varying link traffic counts should increase the estimate precision by reducing the time-dependent OD flows' variance. Since the value of information carried by different links is different, it is important to the transportation agencies to deploy sensors on those links that can bring maximal value of information in order to improve the demand estimation quality. Given the deployment and maintenance costs of such installations, most agencies are called upon to determine the number and locations of such sensors across a given network. However, most of the existing OD estimation methods have been proposed and/or implemented under the assumption of fixed link sensor locations.

A number of researchers have addressed limited versions of the sensor location problem. Most of them formulated the sensor location problem as a flow capture and OD coverage problem (Lam and Lo, 1990; Yang and Zhou, 1998; Yim *et. al*, 1998; Bianco *et al.*, 2001). Zhou and List (2006) focused on locating a limited set of traffic counting stations and automatic vehicle identification readers in a network so as to maximize expected information gain for the subsequent origin destination demand estimation problem solution. However, their methods neither took into account the interrelation between the sensor coverage and sensor location, nor applicability in the context of dynamic traffic assignment.

1.1.2 Research objectives

Driven by the aforementioned motivations, this dissertation addresses a series of problems pertaining to deploying finite resources and generating a network detection system in a manner that produces minimal estimation errors and maximal OD coverage for large-scale urban transportation networks under both deterministic and stochastic traffic conditions. To address the above problem, the fundamental objectives of this research include:

1. Formulate and develop a sensor location model that identifies a set of sensor locations to maximize the coverage of origin-destination (OD) demand flows while

minimizing the demand uncertainty in the estimated OD demand matrix of the road network based on dynamic traffic assignment methodology.

2. Extend the deterministic optimal sensor location model and develop a more robust sensor location model that accounts for the demand uncertainty in the dual objectives of maximizing long run average demand coverage and information gain.

3. Develop and test efficient algorithmic implementations for the proposed sensor location models to find the optimal/near optimal solution for this $\mathcal{NP}$-hard problem with respect to deterministic and stochastic scenarios.

4. Develop an effective framework for clarifying the value of information brought by additional measurement as well as the interactions among different sensors.

The first objective is mainly intended to optimize sensor numbers and locations in the context of known time-dependent OD demand matrices in order to maximize the coverage of demand flows and minimize the demand uncertainty. Due to the day-to-day traffic pattern evolution, the *a priori*/historical demand table used in the estimation problem formulation (described in the next chapter) may be out-of-date and not reflect the prevailing dynamic traffic pattern. It would not be appropriate then to load those demands into the network as part of the procedure for finding optimal sensor locations. As a matter of fact, an up-to-date origin-destination (OD) matrix is imperative in order to find robust sensor locations and sensor coverage that can accommodate network disruptions and other uncertainties, such as special events and weather.

In order to characterize those factors, the demand is viewed as a linear combination of regular demand, demand structural deviation and random dispersion

(Mahmassani and Zhou, 2005). The regular demand is given by the *a priori* OD demand table that can be obtained from survey methods. The structural deviation of real-time demand from daily traffic pattern is used to accommodate the network uncertainties. The random dispersion reflects the other unobserved/unquantifiable factors of the network as well as the inherent stochastic of the daily demand.

The assignment matrix maps the OD flows onto the link counts and is itself dependent on the unknown time-dependent demand flows. It captures three aspects of a traffic network: the network topology, the route choice model and the travel time across the network (Bierlaire and Crittin 2004). Consequently, it plays a critical role in the sensor location problem. It has been a challenging and important work to model time-dependent assignment matrices.

Two scenarios are taken into account in formulating the dynamic sensor location problem. In the first scenario, the minimal optimal sensor locations are exploited under the assumption of no budgetary constraint. The second scenario depicts the more general and practical situation where the transportation agency look for a sensor location plan to deploy finite sensors in large-scale urban transportation networks. To reveal the interrelations among sensor locations, sensor coverage, unknown actual OD demand and traffic assignment, there is a great need to explore ways to allocate sensors so as to generate a network detection system in a manner that produces minimal estimation errors at the minimal equipment cost.

The second objective is to provide a sensor location model when traffic dynamics and network uncertainty are accounted for. Although the traffic dynamics are considered under the first objective, the traffic network is assumed under deterministic conditions. The second objective is intended to extend the dynamic sensor location model under the assumption of recurrent traffic conditions, and incorporate the network uncertainty in a mathematic formulation. As part of a network planning problem, transportation agencies and planners have to deploy limited sensors in the network before the occureence of any non-anticipatable events (e.g. incidents, weather, special events, etc). However, due to the unavoidable randomly occurring uncertain events which consequently affect the traffic pattern in the traffic network, there is a great need to propose a methodology to identify a robust sensor location strategy, which is less sensitive to the network uncertainties.

The third objective is to build an efficient algorithmic procedure specific to the proposed sensor location models. The major concern for the algorithm is to be able to find the optimal or near optimal solutions for the proposed problem with sufficient accuracy and computational tractability. Due to the nature of the combinatorial optimization problem, it is difficult to exhaustively explore the feasible region and make discrete choices. In reality, this area of discrete mathematics is of practical use and has attracted much attention over the years. Constructing an algorithmic procedure for the proposed sensor location models under deterministic and stochastic scenarios such that it can find near-optimal solutions within reasonable running time is imperative. In addition, the flexibility and ease of implementation of the solution algorithm must be taken into account in order to successfully handle different real-world applications.

The fourth objective is trying to illuminate the contribution of the marginal value from additional measurement. A sensitivity analysis based method is essential to apply on a real-world traffic network to study the degree of the demand estimation error correction influenced by different levels of detection and different sensor locations in a portion of a realistic network. This analysis will provide valuable insights about the process of selecting the informative locations for sensors in a network. More interestingly, the comparison between the sensitivity analysis results and the output from the proposed sensor location model can be used to validate the quality and effectiveness of the proposed methodology.

In sum, the overall objective of this dissertation is to build a framework that can help transportation agencies and planners to determine the non-dominated sensor location solutions in terms of maximizing OD coverage and information gains for real-time traffic estimation and prediction in large-scale networks. In addition, a flexible and easily implemented algorithmic procedure is essential in the proposed methodology to the actual large urban transportation networks applications.

1.2 Overview of approach

The conceptual framework presented in this dissertation interprets the sensor location problem as a value of information problem, which leads to interpretation with learning process models. This dissertation aims to present a robust sensor location model to enhance the network state estimate and prediction quality and reduce the uncertainty of estimated OD demand matrices under various network conditions.

Given historical demand and link observation data, this research starts from an objective of minimizing the deviation between the observed and historical link flow counts with a general least square estimator (GLS). In order to accommodate network disruptions, a structural state space model (Zhou 2004) is used to represent the actual demand, which is decomposed into three components:

actual demand= historical demand +structural deviation +random dispersion

The deviation forms are used in this research since they could capture the dynamic traffic pattern temporally and spatially. Moreover, the deviation between actual demand and historical demand subsumes the day-to-day evolutionary information. Consequently, the structural deviation is modeled as the state variable and the objective is to minimize the random dispersion with the GLS estimator.

As the new observation data become available during each observation time interval, the *a priori* historical demand table can be correspondingly updated. Since the value of information obtained from various links is not the same, the problem of concern in this dissertation becomes to find informative sensor locations such that the uncertainties of the dynamic demand inputs are minimized. As an incremental algorithm, a Kalman filter algorithm is a well known approach that can be used to solve a least squares problem in a real-time context. In this dissertation, a Kalman filtering based bi-objective model is formulated to improve demand estimation quality and maximize the OD coverage.

An intuitive thought for solving the proposed model is selecting n_m links every time from the directed network $G(V, A)$, calculating the total link gains each time and selecting the locations having the largest link gains. However the combination of n_m links from total n_{LK} links is $\binom{n_{LK}}{n_m} = \frac{n_{LK}!}{n_m!(n_{LK} - n_m)!}$ which will result in a non-polynomial computational time. Geoffrion (1970) developed a conceptual framework that helps to categorize the methods and solution strategies for large-scale mathematical programming. He called the first category as “master” problems which include Projection, Inner Linearization and Outer Linearization. The second category consists of solution strategies that can be used to solve the master problems in the first category, which include Piecewise, Restriction and Relaxation. However, applying those exact algorithms to the proposed sensor location problem would consume greater computational resources and require additional attention to different realizations.

A DTA simulation-based bi-level programming technique is used to solve the proposed model. In the upper level, a hybrid Greedy Randomized Adaptive Search Procedure (HGRASP), which is a combinatorial optimization algorithm, is developed to find the feasible solution through reducing the effective size of the feasible solution space and exploring the space efficiently. In the lower level, the selected locations from the upper level are evaluated using the simulated results, e.g. assignment matrix, link information gains, etc. through running user equilibrium (UE) of DYNASMART-P (Peeta and Mahmassani (1995). As a dynamic traffic assignment (DTA) based simulation tool, DYNASMART-P is used to propagate vehicles along their prescribed paths and determine the network traffic state. The information about the simulation package can be found in next chapter.

When such improvements are being made on the sensor location problem, a natural extension of the dynamic sensor location model is to account for the network uncertainty directly into the model formulation. Uncertainty is one of the important factors that transportation planners and decision makers have to contend with in making sensor deployment decisions into a traffic network. The high uncertainties, such as locations, durations, severities, induced by the most disasters cause the deterministic sensor location model to be less relevant and the nature of this planning problem makes itself to a two-stage sequence of decisions. The first stage decision denotes a strategic sensor location plan before observations of any randomness events, while the recourse function associated with the second stage denotes the expected cost of taking corrective actions to the first stage solution after the occurrence of the random events. Thus, the proposed sensor location problem is further formulated as a two-stage stochastic model with recourse under network uncertainty in this research. One important view of the stochastic problem is *nonanticipativity*, which means the planning decisions must be made before a random event is observed. In other words, the planning decision is made while the random variables are still unknown, so the decision cannot be determined based on any particular realized values of the random variables. By viewing the sensor location problem as a stochastic optimization problem that takes the network uncertainty into account, the aim of the model is to determine robust sensor locations that may not be optimal to every possible realization of the un-anticipatory scenarios, but will provide good performance under any scenario and perform more robustly with regard to extreme cases. A modified dynamic traffic assignment (DTA)-based HGRASP solution procedure is proposed in conjunction with an incident generation model to find the optimal sensor location plan.

Numerical examples on realistic networks are used to illustrate the proposed models and solution algorithms. In addition, a sensitivity analysis is conducted to systematically evaluate different sets of sensor locations under certain criteria, such as adjacency rules in a large-scale urban transportation network. The analysis considers both randomly generated location scenarios as well as scenarios based on engineering judgment. The latter considers placing sensors on high volume links on the main freeways and arterials. Taken together, the two sets of scenarios provide useful insight into the robustness of the real-time DTA estimation and prediction, and the effect of location-specific considerations on estimation and prediction quality.

The test results indicate that the solution of the proposed model is consistent and robust under different traffic conditions.

1.3 Dissertation organization

This dissertation comprises six chapters. The second chapter provides an overview and discussion of several topics, including OD demand estimation and prediction, sensor location problem, and previous related stochastic network design research. It also briefly introduces the DYNASMART simulation package. Chapter 3 first presents a conceptual Kalman filtering based framework for the sensor location problem, and a theoretical description of the goals associated with the sensor location problem. Then a bi-objective model is proposed followed by the Hybrid Greedy Randomized Adaptive Search Procedure (HGRASP) algorithmic procedure. Numerical examples are used to illustrate the proposed methodology. Taking the network uncertainty into account, Chapter 4 extends the deterministic optimal sensor location problem proposed in Chapter 3 to a stochastic optimal sensor location problem and presents a modified HGRASP-DTA solution procedure in conjunction with an incident generation model. Chapter 5 includes an analysis that illustrates how estimation and prediction of the network performance can be influenced by the location and number of detectors in the network. Sensitivity analysis and the proposed bi-objective model are applied to implement a series of experiments on a real-world large-scale urban transportation network. Chapter 6 concludes this research and delineates some possible areas for further research.

Chapter 2 Background Review

2.1 Introduction

Origin-destination (OD) demand matrices play a critical role in many important transportation research problems from traffic operation control to transportation network planning analysis. As an input to many transportation applications, an accurate OD demand matrix becomes extremely important because the link flows after loading the demand matrix must be close to the actual values in order to estimate the network state conditions. High cost in terms of time, budget, manpower, etc of traditional methods that combine household-based interviews and roadside surveys limit the usefulness of this method in many applications, especially in the context of real-time traffic estimation and prediction.

Information technologies have great potential in improving the network state estimation and prediction quality. Recent advances in wireless networking and sensor networks significantly have impacted the design of intelligent transportation systems (ITS) to make transportation systems safer and more efficient. Numerous exciting research challenges exist for designing wireless networking and sensor network technologies for vehicle to vehicle, vehicle to infrastructure, and within infrastructure sensing and communication applications. Due to the relatively low cost and ease of obtaining network sensor data, many studies have been conducted regarding the methods for analyzing the collected link traffic data to estimate or predict OD demand matrices. However, most of those studies were implemented on the assumption of fixed sensor locations in the network. Understanding the relationship between sensor location and the quality of the estimated OD demand, as well as trade-offs between sensor investments and information gain are critical to the agencies' decision-making in this regard. A number of researchers have addressed limited versions of the sensor location problem.

In the following sections, the relevant studies are reviewed. First, the off-line and on-line time-dependent OD estimation and prediction methods are described in section 2.2. Section 2.3 reports study efforts on sensor location over the past three decades. Section 2.4 reviews the literature on stochastic programming approaches, then a simulation-based dynamic traffic assignment system DYNASMART is introduced. Finally, the main conclusions are summarized in the closing section.

2.2 Overview of methods for estimating O-D matrices

2.2.1 Methods for off-line O-D estimation

Due to the day-to-day traffic pattern evolution, an up-to-date origin-destination (OD) matrix is important for real-time network traffic estimation and prediction, which integrates the a priori matrix with link counts obtained from the low-cost road side sensor stations. The past three decades have seen many studies on OD matrix estimation. In general, those studies can be grouped in two categories, traffic assignment based approaches and statistical inference approaches.

The first category includes "information minimization" (entropy maximization) models. Zuylen and Williumsen (1980) developed two models based on information minimization and entropy maximization principles to estimate an OD matrix from traffic counts by reproducing the observed link flows. They introduced a variable p_{ij}^{a} to represent the proportion from origin i to destination j that use link a and assumed proportional traffic assignment. Although they introduced minimum external information into the model and turned the problem into a multi-proportional problem, the assumption that the assignment matrix is independent of the OD flow limited the applicability of their procedure in real world congested networks. Fisk (1988) took into account the congestion factor that impacts travel times and consequently influences drivers' path choices and assignment matrix. She combined the maximum entropy model and user-equilibrium model into a single mathematical problem and transformed the problem into a bi-level programming formulation.

Recognizing that the OD estimation problem is usually an under-specified problem, in that the number of OD pairs, which are the unknown variables in this problem, is normally greater than the number of link traffic stations, researchers integrated the *a priori* OD matrix with the link counts in order to obtain a unique estimated OD matrix. The second category includes maximum likelihood (ML) approach, generalized least squares (GLS) approaches and Bayesian Inference approach. Spiess (1987) assumed the OD demand can be obtained from independent Poisson distributed random variables with unknown means. A ML model was formulated to estimate these means to reproduce the estimated link flows consistent with the observed link flows. However, his study assumed the assignment matrix is constant and determined exogenously. Cascetta (1984) proposed a generalized least squares estimator that combines traffic counts with an assignment model. Bell (1991) incorporated the inequality constraints and presented a simple iterative algorithm for solving the constrained GLS problem and proved its convergence. Maher (1983) assumed the *priori* OD matrix and the observed link counts follow multivariate normal distributions and proposed a Bayesian statistical inference based model to update the *priori* OD matrix. The general formulation of the static OD estimation is as follows

$$J = \arg\min\left[(C - H\hat{D})^T R^{-1}(C - H\hat{D}) + (D - \hat{D})^T V^{-1}(D - \hat{D})\right]$$
$$s.t. \quad \hat{D} \geq 0$$

where R and V are dispersion matrices, D and $\hat{D}$ are the target and estimated demand matrix, respectively.

Most of the static OD estimation methods assumed that the assignment matrix is constant (proportional assignment) and independent of the OD flows. The earliest reported study to estimate "time-dependent" OD matrices was implemented for dynamic origin-destination flows estimation in an interchange or corridor (Cremer, *et. al* 1981). Cascetta *et al.* (1993) extended and generalized the static OD estimation model and proposed two approaches, simultaneous and sequential estimators, to estimate dynamic OD matrices by dynamic traffic assignment modeling. The simultaneous approach estimates the entire OD demand pattern by using counts over all intervals simultaneously.

$$\mathbf{d} = \arg\min_{s_1 \geq 0,..s_n \geq 0}\left[f_1(s_1,......,s_n;\hat{d}_1,....,\hat{d}_n) + f_2(s_1,......,s_n;\hat{v}_1,....,\hat{v}_n)\right]$$

In the sequential approach, the demand vectors for a single interval are estimated sequentially.

$$\mathbf{d} = \arg\min_{s_h \geq 0}\left[f_1(s_h,\hat{d}_h) + f_2(\hat{v}_h, s_h;\hat{d}_1,\hat{d}_{h-1})\right]$$

where $f_1(\bullet), f_2(\bullet)$ are the objective functions depending on the distributional assumptions made on the vectors $\hat{d}_h$.

The models for dynamic OD estimation can be categorized into two classes: non-DTA based and DTA-Based. In the non-DTA based class, Wu and Chang (1996) extend Bell's (1991) linear system models and proposed a non-assignment based O-D estimation method with the inclusion of screenlines to estimate time-dependent O-D demand matrices for closed networks. In the DTA based class, Tavana and Mahmassani (2000) proposed a bi-level least-squares estimation method using a dynamic traffic assignment (DTA) based simulation program to estimate time-dependent OD. Zhou, *et. al*.(2003) extended Tavana's model to a bi-objective model with weight function by incorporating *a priori* OD demand table and multi-day link flow counts.

In recent years, with the availability of new technologies for vehicle tracking, automatic vehicle identification (AVI) data have been used to estimate the OD matrix with point sensor data (Van der Zijpp *et al.* (1980), Dixon et al. (2002), Zhou and Mahmassani (2006)).

2.2.2 Methods for real-time dynamic O-D estimation and prediction

Dynamic OD demand estimation and prediction is a critical component for real-time dynamic traffic assignment. As unknown variable, the time-dependent OD demand involves both temporal and spatial dimensions. With respect to the OD demand, real-time OD estimation and prediction has become an important element in dynamic traffic management systems (Ashok, et. al 1993).

The basic problem of OD prediction is to compute, in real-time, the future OD estimates with current network traffic information, such as link counts and proportions in conjunction with historical OD flows. Several approaches have been proposed in the literature to model the dynamic nature of demand. Okutani (1987) proposed a state-space model using an autoregressive process on the OD flows as the transition equation to capture temporal interdependencies. However he ignored the pattern of the OD trips in a transportation network is determined not only by the spatial but temporal distribution of traffic activities, which cannot be modeled by a simplistic auto-regressive process. By recognizing this limitation, Ashok *et al.* (1993) used deviations of O-D flows from best historical estimates instead of the O-D flows themselves as state-vector in a state-space model. Because of the estimations of not only current interval, but prior intervals state variables, his method is very computationally intensive. Kachroo, Narayanan and Ozbay (1995) extended this approach to account for colored noise in the system. Based on their previous work, Ashok and Ben-Akiva (2000) proposed two approaches for real-time estimation and prediction of time-dependent OD flow. The first approach is an extension of the autoregressive model using the deviation between the actual and historical OD flows. In order to keep the estimation procedure computationally tractable, they used augmented state-vector and assumed the OD flows in prior time interval hold constant. They used the deviations of departure rate from each origin and shares headed to each destination in the second approach.

Recognizing the fact that the prediction of OD matrices and other network traffic conditions is more reliable in the near-term (roll period), Peeta and

Mahmassani (1995) proposed a rolling approach, previously used in the production-inventory control literature, to solve large-scale network dynamic assignment in quasi-real-time. The rolling horizon implementation of the DTA model recognizes that prediction of OD matrices and network conditions is more accurate in the short term (roll period), while the uncertainty increases beyond this period. Rather than assuming that time-dependent OD matrices and network conditions are known a *priori* for the entire assignment horizon, a more realistic scenario is to consider that the information of short-period dynamic OD matrices and network conditions is deterministic, whereas information beyond this short period (roll period) will not be available until some time later. To illustrate their approach, figure 2-1 shows two consecutive stages of PDYNA as well as the interrelationship between PDYNA and OD estimation. The stage length of a PYDNA is h units and the simulated link proportions in this stage (stage σ-1) and real-time traffic measurements are provided to OD demand estimation module for the OD estimation calculation. Following the OD demand estimation, the OD prediction component predicts the OD demands of the future time period η on the basis of current OD estimation results. The predicted OD demand will be utilized by the next PDYNA for predicting network traffic flow propagation in stage σ. To guarantee that PDYNA in stage σ finds the OD information it is requesting, the prediction horizon η has to be greater than h. Similarly, to guarantee that OD estimation in stage k+1 receives the predicted link proportions, the OD estimation state length γ must be less than the PDYNA stage length h minus the roll period *l*. The shaded portion of stage k represents the short-term duration for which demand information is consider reliable and is referred to as the roll period of l time units. Beyond this point, the OD forecasting and other network conditions in the rest port of the stage k are subject to substantial uncertainty.

The Kalman filter algorithm has been used to accommodate the requirements of real-time OD estimation and prediction (Okutani 1987; Ashok and Ben-Akiva 1993, 2000; Wu and Chang 1996; Kang 1999; Zhou 2004). This algorithm is a recursive method that gives a linear, unbiased, and minimum error variance estimate of the unknown state vector at each time instant with the incoming observation data. Inspired by Ashok and Ben-Akiva (1993)'s work, Bierlaire and Crittin (2004) derived a least-square model and used the LSQR algorithm to overcome Kalman filter algorithm's inability to handle large-scale network. Wu (1997) proposed a revised Multiplicative Algebraic Reconstruction Technology (MART) algorithm based on a normalization treatment and the diagonal searching technique from the nonlinear programming methodology for online OD flow updating.

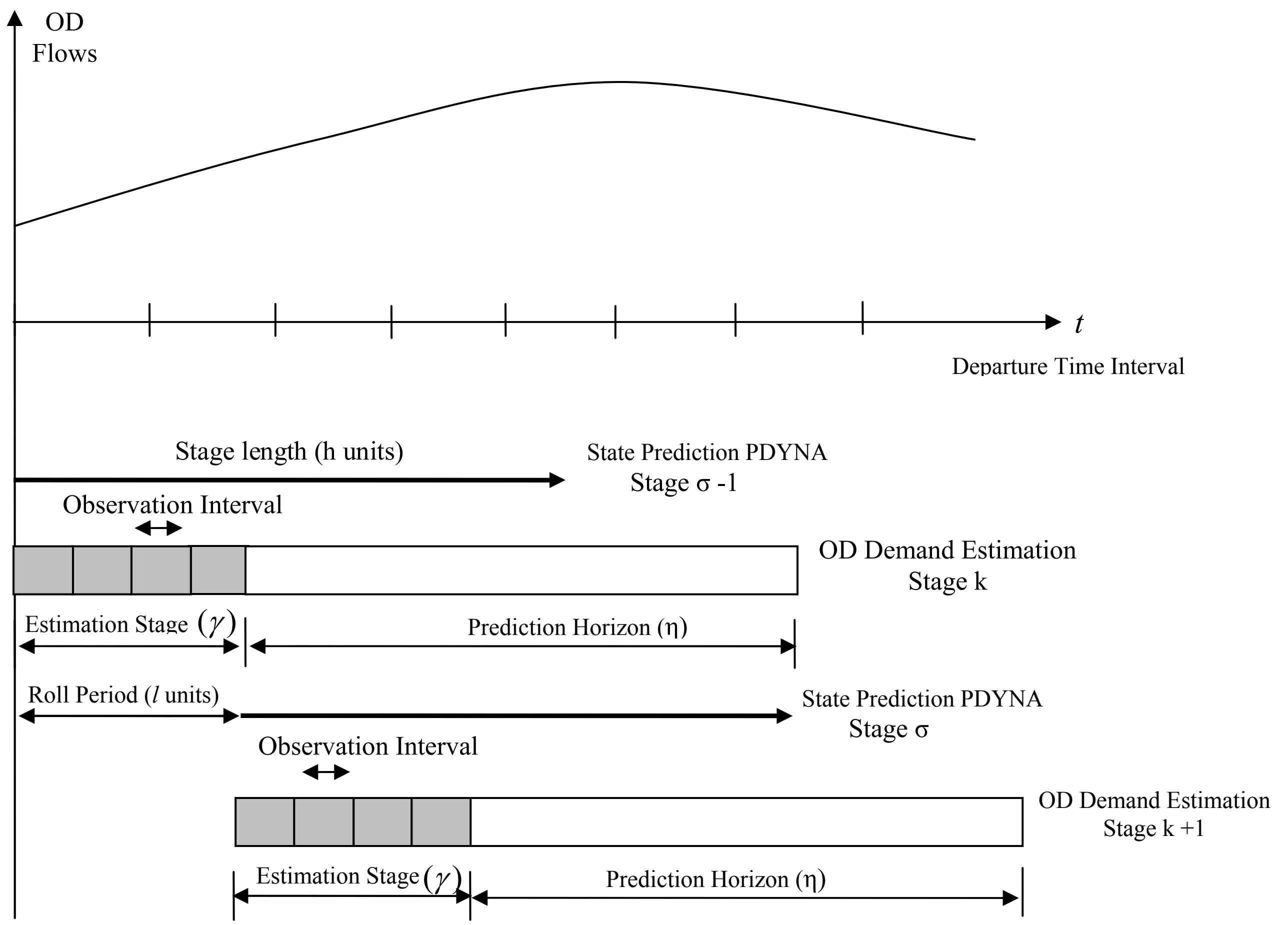


Figure 2-1 Rolling Horizon Solution

In order to capture the dynamic nature and nonlinear trend characteristics, Mahmassani *et .al.* (1998) and Kang (1999) introduced a general polynomial transformation framework to formulate the dynamic OD estimation and prediction and combined it with a Kalman Filter model under the assumption that OD flows will not dramatically change within an estimation period. The polynomial trend filter equation can be described as follows

$$\mathbf{d}_{k+\tau|k} = \sum_{p=0}^{h} \varphi_p \tau^p = \varphi_0 + \varphi_1 \tau + ... + \varphi_h \tau^h$$

where h is the order term, and φ_p is the polynomial coefficient vector.

By using the transformation, the OD estimation and prediction problem becomes an over-determined problem and the state variables domain becomes linear or quasi-linear which is the requirement for getting optimal estimation results using the Kalman Filtering method. Combining Ashok *et. al* (1993) and Kang's model, Zhou (2004) proposed a structural state space model, a Kalman Filtering based OD estimation and prediction model, which can be integrated into a DTA simulation framework. He integrated historical demand information as well as structural changes into a real-time demand process model, in order to provide accurate and robust demand prediction under recurrent and non-recurrent conditions. He proposed a linear model combing *a priori* OD estimate $\hat{d}^{\tau}_{(i,j)}$, structure deviation $\chi^{\tau}_{(i,j)}$ and random disturbance $\varepsilon^{\tau}_{(i,j)}$ together. By integrating the regular demand pattern, his structure model leads to smaller estimation and prediction variance compared to a pure polynomial model. The recursive dynamic OD demand estimation and prediction procedure integrating the structural state space model and Bang-Bang control logic for the real-time traffic system is described as follows.

<u>Real-time dynamic O-D estimation and prediction:</u>

Step 0: Initialization

Let the initial estimation value be $\hat{X}_{0|0} = E(X_0)$, $P_{0|0} = Var(X_0)$, $k = 1$

Step 1: Prediction

Predict the mean and covariance estimates from state $k-1$ to state k after using measurements obtained at state $k-1$ and correcting the state variable estimates at state $k-1$

$$\hat{X}_{k|k-1} = A_{k-1}\hat{X}_{k-1|k-1}$$

$$P_{k|k-1} = A_{k-1}P_{k-1|k-1}A_{k-1}^{T} + W_{k-1}$$

where $A_{k-1} = Diag(A_{k-1}^{1}, A_{k-1}^{2}, ...A_{k-1}^{j}...., A_{k-1}^{N_{od}})$, $A_{k-1}^{j} = \begin{bmatrix} 1 & \beta & \frac{\beta^2}{2!} \\ & 1 & \beta \\ & & 1 \end{bmatrix}$

Step 2: Estimation and Correction

After obtaining the new link proportions and link observation data, the Kalman filter gain matrix at state k is calculated as follows,

$$K_k = P_{k|k-1}H_k^T(H_kP_{k|k-1}H_k^T + R_k)^{-1}$$

Using K_k correcting the predicted $\hat{X}_{k|k-1}$ and $P_{k|k-1}$ with the link observation data.

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k(Y_k - H_k\hat{X}_{k|k-1})$$

$$P_{k|k} = (I - K_kH_k)P_{k|k-1}$$

Step 3: Demand Deviation Error Checking

If $\hat{X}_{k|k} \in [L,\ U]$, go to step 4, Otherwise, if $\hat{X}_{k|k} < L$, Let $\hat{X}_{k|k} = L$. If $\hat{X}_{k|k} > U$, Let $\hat{X}_{k|k} = U$, Where L and U are the lower boundary and upper boundary of the demand deviations.

Step 4: Estimation of real-time demand

After obtaining the new demand deviations $\hat{X}_{k|k}$, update the *a priori* estimated OD demand using $\hat{X}_{k|k}$. From (1), it can get

$$E(d^{\tau}_{(i,j)}) = E(\hat{d}^{\tau}_{(i,j)} + \chi^{\tau}_{(i,j)} + \xi^{\tau}_{(i,j)})$$
$$= \hat{d}^{\tau}_{(i,j)} + X^{\tau}_{i,j}$$

Step 5: If reaches the simulation horizon, done; otherwise let $k = k + 1$, go to step 1.

2.2.3 Methods for distributed dynamic O-D demand

As noted, the focus for application of dynamic traffic assignment models to support real-time traffic management decisions requires the ability to execute these procedures on large-scale, real-world networks. As such, it is not sufficient to develop and illustrate procedures that may work on a small network, in order to demonstrate algorithm design issues or properties. It is necessary to address the challenges associated with real-world networks and applications. A major challenge arises from the need to process large amounts of traffic data and generate information supply strategies in real-time, resulting in computationally intensive control architectures that are often a key barrier to their implementation. Building a dynamic O-D distributed modeling framework is a logical approach to overcome the limitations of current-generation computing platforms.

Generally, decomposition approaches applied in the DTA arena can be classified into three categories: (1) distributing independent work onto different CPU's; Peeta, et al (1999, 2004) distributed the system optimization and user equilibrium of the Multiple User Classes Time-Dependent Traffic Assignment (MUCTCDTA) algorithm onto different computers; (2) developing more computationally efficient algorithms for parallel/distributed modes (Ziliaskopoulos, *et al* 1997, Jiang 2004, Lo *et al* 1999); and (3) light global control/independent subnets design (Hawas *et al* 1997, Jayakrishnan *et al* 1999, Chiu, *et al* 2002, Liu *et al* 2004). Decomposition approaches of network OD demand matrices for large scale networks have gained considerable attention as a research topic that is attracting larger numbers of researchers working in this field.

2.3 Overview of Sensor Location Problem Approaches

O-D demand estimation using link traffic counts is a well known underspecified problem, in that the number of links with count stations is usually less than the total number of O-D pairs in the network. As a matter of fact, not all of the links convey the same amount of information; some links even make no or slight contribution to update/improve the *a priori* OD matrix. Thus, how to deploy a limited number of sensors in a traffic network to achieve maximal information content in the observed data and increase the reliability of an estimated O-D matrix becomes an important research topic.

Although the quality and quantity of sensor data are considered as essential inputs to an OD estimation problem, most of the demand estimation and prediction methods were built under the assumption of a given subset of link sensors. Aware of the inherent connection between the OD estimation and link observation counts, several researchers have approached the sensor location problem as an OD covering problem. Lam and Lo (1990) proposed "traffic flow volume" and "O-D coverage" criteria to determine priorities for locating sensors. By employing a concept of maximum possible relative error (MPRE) to bound the real relative error, Yang *et al.* (1991) formulated a simple quadratic programming problem and showed that if an

OD pair is not covered by a sensor, the MPRE is infinite. The MPRE is defined as the maximum possible relative deviation of the estimated OD matrix from the true one

$$MPRE(\theta) = \max \sqrt{\sum_{w \in W} \theta_w^2 / |W|}$$

$$\text{Subject to: } \sum_{w \in W} p_{aw} t_w \theta_w = 0$$

where $\theta = \dfrac{d_w^* - d_w}{d_w}$, $|W|$ is the number of OD pairs, and d_w^* is the actual OD flow.

Yang and Zhou (1998) further proposed four basic rules for the sensor location problem based on the MPRE.

- Rule 1: *OD covering rule*: A certain portion of trips between any OD pair should be observed.
- Rule 2: *Maximal flow fraction rule*: For a particular OD pair, link with the maximal fraction of that OD flow should be selected.
- Rule 3: *Maximal flow-intercepting rule*: Under a certain number of sensor constraint, the maximal OD pairs should be observed.
- Rule 4*: Link independent rule*: The resultant traffic counts on the selected links should not be linearly dependent.

Ehlert *et al*.(2006) extended Yang and Zhou's work by taking the existing sensors into account and sought the second-best solution. Yim *et al.* (1998) evaluated maximal net O-D capture rule and maximal total O-D captured rule on a large-scale network. Bianco *et al.* (2001) proposed an iterative two-stage procedure which focuses on maximizing "coverage" in terms of geographical connectivity and size of the O-D demand population. Chootinan *et al.* (2005) formulated a bi-objective model to locate traffic counting stations for the purpose of OD matrix estimation. They considered the maximal covering rule while minimizing the sensor quantity as two conflict conditions and proposed a multi-objective method to balance those two criteria. Yang *et al.* (2006) formulated an integer linear programming model to solve a screen line-based sensor location problem. Pravinvongvuth *et al*. (2005) proposed a methodology for selecting the preferred plan from the set of Pareto optimal solutions obtained form solving the multi-objective automatic vehicle identification (AVI) reader location problem constrained by the resource limitation as well as the O-D flow coverage.

Based on the assumptions that an active sensor can provide path flows and each edge in the network associated with exact two paths, Gentili and Mirachandani (2005) considered the sensor location problem as a set covering problem and proposed some graph theoretic based models to locate *active* path-ID sensors on a network. They presented a problem formulation and analyzed three different scenarios depending on the number of conventional (passive) sensors already installed in the network. However, they did not take into account the factors that link volume and correlations among different sensors may also influence the sensor locations in a network. Moreover, their assumptions that tried to capture all of the network path flows in conjunction with one link associated exactly with two path flows may be difficult to fulfill in terms of the market penetration rate and the uncertainty of the travelers' route choices decisions due to the anticipated or un-anticipated network traffic disturbances in a general road network, especially in a large-scale congested network.

The general approach used to address this problem relies on heuristics, especially greedy algorithms (Yang and Zhou 1998). These algorithms basically seek to find the most important location first and locate a sensor there. Then find second-most important location and continue until reaching a pre-specified termination criterion (# of sensors or no significant improvement).

The aforementioned studies were all implemented under the measurement error free assumption, and their objective is maximization of O-D coverage. None of the studies were intended to reduce the uncertainty in the O-D matrix estimation through sensor deployment. Zhou and List (2006) focused on locating a limited set of traffic counting stations and automatic vehicle identification readers in a network so as to maximize expected information gain for the subsequent origin destination demand estimation problem solution.

All existing sensor location approaches assume that static traffic patterns on the network prevail. Those methodologies ignored an import common source of temporal variability in the link-level performance, the nonstationary characteristics of cross-traffic, which leads to the static models unable to capture the traffic dynamics. In addition, the static sensor location models are not robust under different traffic conditions.

2.4 Overview of Stochastic Programming Approaches and Incident Generation Approaches

The transportation system is one of the most complicated dynamic social systems, as it includes road systems, vehicles, control systems as well as the inherent uncertainties due to the interactions among different components or unavoidable unpredictability (randomness) caused by disasters, such as hurricane, earthquake, flood, bio/chemical/nuclear hazards or traffic incidents.

Mahmassani (1984) presented an overview of evaluation approaches for uncertainty in transportation systems. He categorized five different types of uncertainties in the evaluation of transportation systems. (1)Unexpected events and unforeseen situations, such as major political disturbances or unanticipated technological fails; (2) The exogenous states affecting the transportation systems, such as new administration, economic boom or bust etc; (3) Uncertainty in the values of measured or predicted impacts usually as a result of the modeling activity; (4) Fuzziness or vagueness characterized with the description of a performance measure in transportation systems; (5) Uncertainty as to the preferential or normative basis of the evaluation. This includes inclusion uncertainty, appropriate trade-offs among criteria, the risk attitudes of the decision makers in the decision process, the biases of the actors in the planning process. The approaches to deal with those uncertainties include (1) Reducing uncertainty; (2) Structuring the decision process; (3) Evaluation and design criteria and guidelines; and (4) Explicit evaluation techniques.

A stochastic programming model can incorporate the uncertainties into the formulation. Two types of models are usually studied: (i) Multi-stage recourse problems and (ii) Chance constrained problems. A traditional two-stage stochastic programming with recourse model is formulated into two stages. Decisions are implemented before the random events are observed in the first stage, after which, a response action made in the second stage is applied to each outcome of the random events that might be observed in the first stage.

The classical two-stage stochastic linear program model with recourse was first proposed by Dantzig (1955) and Beale (1955) to solve the linear model under uncertainty, which can be formulated as follows (Birge 1997):

$$
\begin{aligned}
& MinZ = c^T x + E_{\zeta}\left[\min q(\omega)^T y(\omega)\right] \\
& s.t.\ \ Ax = b \\
& \quad T(\omega)x + Wy(\omega) = h(\omega), \\
& \quad x \geq 0, y(\omega) \geq 0
\end{aligned}
$$

where x, y are variables, c, b, A are parameters, q, T, h are realization-dependent random variables for each ω. $\omega \in \mathbf{\Omega}$ denotes the system realization of random events; For a given realization ω, the second stage problem data $q(\omega), T(\omega), h(\omega)$ become known. If the recourse function in the second stage is given, the stochastic program can be converted to an ordinary deterministic equivalent program.

Stochastic mathematical models have been widely applied in the transportation and operation research areas. Gendreau et al. (1996) reviewed the stochastic vehicle routing studies during the past decades from a theoretical aspect. Waller and Ziliaskopoulos (2001) introduced a two stage stochastic model with recourse to solve the network design problem by accounting for uncertain network system demand and traffic conditions. Sawaya et al. (2001) proposed a multistage stochastic model with recourse to design real-time traffic control strategies to respond the freeway congestion caused by unexpected incidents through taking into account demand variations and incident severities. Liu and Fan (2007) introduced a two stage stochastic model to support making retrofit decisions with considering the random occurred earthquakes.

Uncertainty in demand may result in the underestimation of the system performance, such as total system travel time, which leads to sub-optimal planning decisions (Waller, Schofer and Ziliaskopoulos, 2001). The potential advantages achieved by explicitly including the minimization of the variation of the estimated OD demand into the objective function include:

1) It will potentially reduce the computation intensiveness and model complexity; in order to develop robust improvement schemes for road network, Waller *et al.*.(2001) analyzed the traffic assignment results by enumerating every possible demand scenario. Yin et al. (2004) proposed sensitivity based model and scenario based model to examine the network travel time under different level of demand. The small range of demand variation resulted from demand uncertainty reduction by strategically deploying sensors in the network, leads to less possible demand scenarios and increases the system robustness;

2) It potentially increases the robustness of the model. A typical stochastic model's objective usually only optimizes the expectation of the distribution of the objective value while ignores the higher moments. Minimization of the expected variance of the estimated OD demand reflects the decision maker's risk aversion to the uncertainty and to find a robust solution that is valid to various possible random scenarios.

A challenge in the sensor location problem is how to detect the occurrence of highly uncertain incidents in the network. The MUTCD (Maryland SHA, 2006) defines a traffic incident as an emergency road user occurrence, a natural disaster, or other unplanned event that affects or impedes the normal flow of traffic. It divides the traffic incidents into three general classes of duration, each of which has unique traffic control characteristics and needs. These classes are: (a). Major—expected duration of more than 2 hours; (b). Intermediate—expected duration of 30 minutes to 2 hours; and (c). Minor—expected duration under 30 minutes.

Martin *et al.* (2001) examined various incident detection technologies, which include computer-based automatic incident Detection (AID), Video Image Processing (VIP) and detection by cellular telephone call-ins. They compared different algorithms, such as pattern recognition, catastrophe theory, statistical, and artificial intelligence, to find the potential incident location. Chiu, *et al.* (2001) assumed the occurrence of n incidents on link a_n follows a Poisson process.. The system uncertainties are conceptually modeled by a scenario tree which describes system uncertainty evolution across all stages.

2.5 Overview of DYNASMART

Dynamic Traffic Assignment (DTA) is a core capability required for the operation of Advanced Transportation Management Systems (ATMS) and Advanced Traveler Information Systems (ATIS). DYNASMART is a state-of-the-art Traffic Estimation and Prediction System (TrEPS) mesoscopic simulation software package for effective support of transportation network planning and operations decisions (offline version DYNASMART-P) and ATMS/ATIS in the ITS environment (real-time online version DYNASMART-X).

2.5.1 Overview of DYNASMART-P

DYNASMART-P is a state-of-the-art dynamic network planning, analysis and evaluation tool. It represents the traffic interactions in the network and models the evolution of traffic flows in a traffic network resulting from the travel decisions of individual drivers. The model is also capable of representing the travel decisions of drivers seeking to fulfill a chain of activities, at different locations in a network, over a given planning horizon. Due to its inherent characteristics that explicitly describe traffic processes and their time-varying properties and explicitly represent traffic network elements, i.e. signal, VMS diversion strategies, etc, DYNASMART-P is more advantageous than static assignment tools.

The embedded components, such as simulation component that moves individual vehicles in the detailed represented network according to macroscopic traffic flow relations under some simulation assignment approach (i.e. SO, UE, one-shot simulation), path-processing component that determines the path level attributes (i.e. travel time) given the link level attributes (i.e. link types, link length, etc.) from the simulator component, behavioral component that provides the drivers in the network alternative paths or additional information (VMS, radio, etc) during non-recurrent congestions, make DYNASMART-P achieve a balance between representation detail, computational efficiency, and input data requirements.

DYNASMART-P generates various performance statistics over time for each link in the network at both the aggregate and disaggregate levels. Those measures of effectiveness (MOE) include vehicle level, such as vehicle trips, speeds, densities and queues, path level, such as vehicle trajectory, and network level, such as average travel times, average stopped times, and the overall number of vehicles in the network.

DYNASMART-P is modeled and featured as an offline operational tool and its primary distinction from the online version (DYNASMART-X, described in 2.5.2) is that DYNASMART-X comprises real-time dynamic traffic assignment descriptive and normative capabilities with other components, such as demand estimation and forecasting, consistency checking and updating, and parallel and distributed capabilities of different mode. With a specific designed data interface, (such as XML, SOAP, Figure 2-4), DYNASMART-X can interact with external real-time sensor data collected throughout the network.

2.5.2 Overview of DYNASMART-X

With widespread deployment of sensor technologies that feed traffic data into modern TMC's, it is imperative to leverage the investment in hardware into tangible benefits for the traveling public. Beyond the traditional responses to traffic incidents, such as police and EMS dispatch, methodological developments such as simulation-based DTA systems contribute to providing real-time decision support capabilities in TMC's. Because they are based on a representation of actual network traffic dynamics, real-time DTA systems enable the estimation and prediction of traffic conditions as events occur and new situations unfold in a network. Predicted information is an important element of next-generation advanced traveler information systems. The ability to evaluate the impact of different operational measures under alternative short-term scenarios is essential to the modern management of transportation corridors.

DYNASMART-X is a state-of-the-art real-time TrEPS (Traffic Estimation and Prediction System) for effective support of Advanced Traffic Management Systems (ATMS) and Advanced Traveler Information Systems (ATIS). DYNASMART-X interacts continuously with multiple sources of real-time information, such as loop detectors, roadside sensors, and vehicle probes, which it integrates with its own model-based representation of the network traffic state. The system combines advanced network algorithms and models of trip-maker behavior in response to information in an assignment-simulation-based framework to provide reliable estimates of network traffic conditions; predictions of network flow patterns over the near and medium terms in response to various contemplated traffic control measures and information dissemination strategies and routing information to guide trip-makers in their travel. One of the most important capabilities of a real-time traffic simulation system, which distinguishes it from a model intended for off-line planning applications, is to be able to estimate and predict time-varying OD demand adaptively with incoming real-time traffic sensor data. Establishing and developing an appropriate OD estimation/prediction model is an essential requirement in DYNASMART-X. State mapping matrices, Kalman filter process noise variance-covariance matrices, and measurement noise variance-covariance matrices are three sets of key parameters required in the current implementation of OD estimation/prediction in DYNASMART-X.

Figure 2-2 illustrates the demand data flow of a real-time DTA system. Based on a Kalman filter real-time OD demand estimation and prediction algorithm, OD estimation module utilizes real-time traffic measurement data (link counts per observation interval) to update OD demand, which followed by the OD demand prediction. By using the predicted OD demand, PDYNA generates simulated link proportions on all observed links and the vehicle routing policy, which will be fetched by OD estimation module for the next several departure intervals and then the OD demand estimation module starts with the variables from last state. Consistency checking and updating is an important function incorporated in DYNASMART-X to ensure consistency of the simulation-assignment model results with actual observations, and to update the estimated state of the system accordingly. Another external support function is intended to perform the estimation and prediction of the origin-destination (OD) trip desires that form the load onto the traffic network, and is as such an essential input to the simulation assignment core.

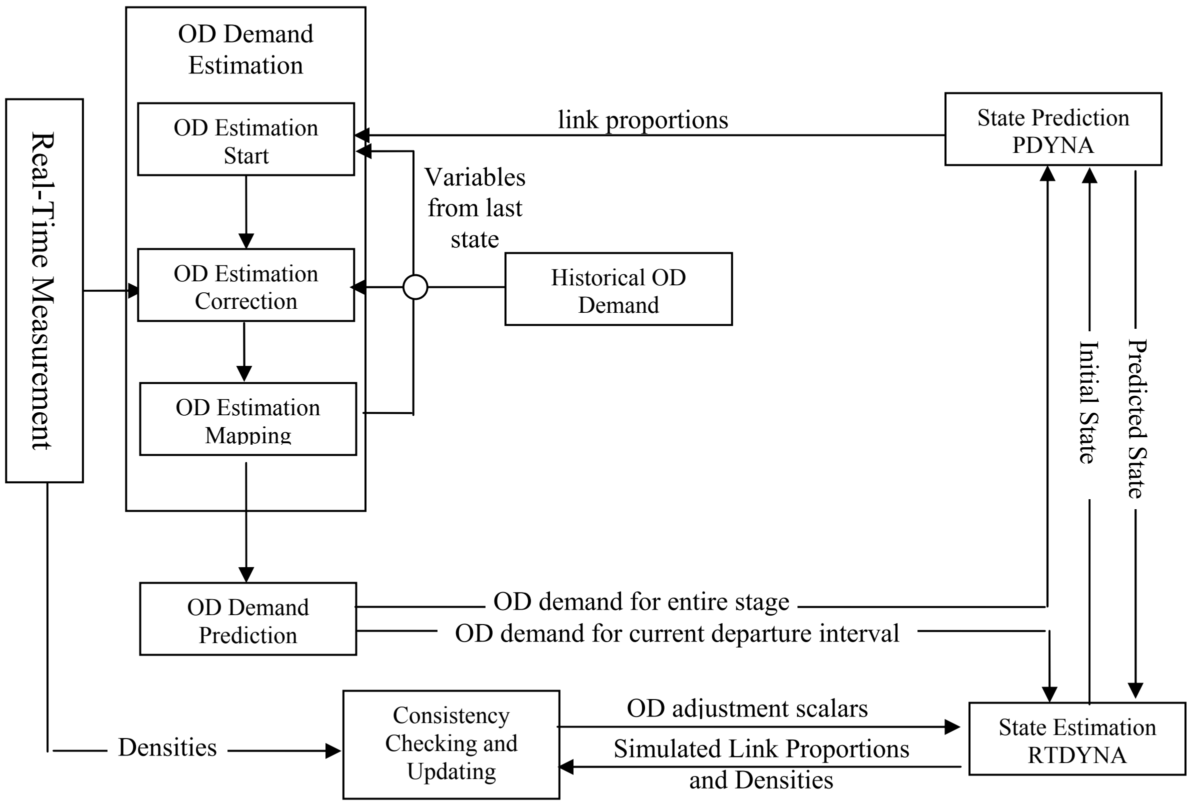


Figure 2-2 Demand Data Flow of Real-Time DTA System

The functionality of DYNASMART-X is achieved through judicious selection of modeling features that achieve a balance between representational detail, computational efficiency and input data requirements (Mahmassani, *et. al* 2002). DYNASMART-X consists of a set of components designed to perform its intended functions. The first component is the graphical user interface (GUI). The second component is the database. The third component comprises the algorithmic modules that perform the DTA functional capabilities. These modules are: 1) state estimation; 2); state prediction; 3) OD estimation; 4) OD prediction; and 5) consistency checking and updating. The fourth and final component is the set of CORBA programs used to implement the scheduler and the data broker. Figure 2.3 depicts a high-level view of the DYNASMART-X system structure and the interrelationship among the components and modules.

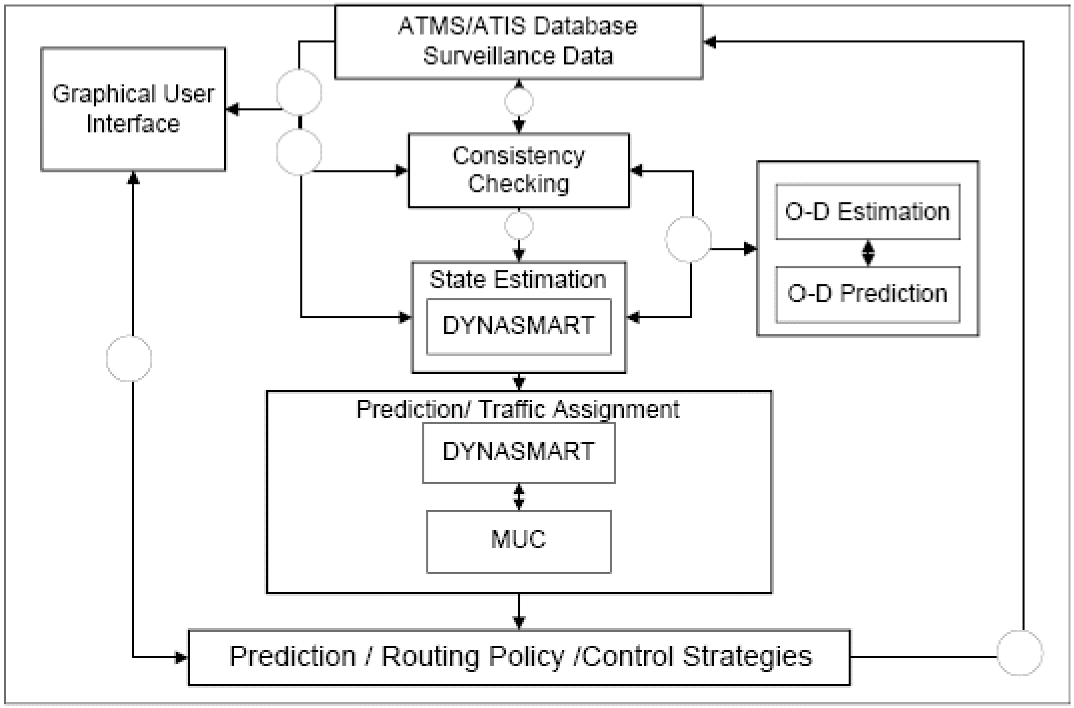


Figure 2-3 DYNASMART-X Functional Diagram

The algorithmic component is the main entity in the system. It is responsible for implementing various DTA tasks. The purpose of the state estimation module (RTDYNA) is to estimate the current traffic states in the network. The state prediction module (PDYNA) on the other hand provides future network traffic states for a pre-defined horizon. The OD estimation module (ODE) is responsible for estimating the coefficients of a time-varying polynomial function that describes the OD demand in the current stage. The OD prediction module (ODP) calculates the demand that is generated from each origin to each destination at each departure time interval of the current and future stages. Finally, the consistency checking modules are responsible for minimizing the deviation or discrepancy between what is estimated by the system and what is occurring in the real world, in an effort to control error propagation. DYNASMART-X implements two levels of consistency checking: short term and long term. The short term one (STCC) uses the link densities and speeds of the simulator to evaluate the consistency of the flow propagation with the real world and correct the simulated speeds. Long term consistency checking (LTCC) calculates scaling factors that are applied in the next execution instance of RTDYNA. An updating function runs in parallel with the STCC and LTCC tasks. The remaining components in the system serve as supporting entities to the algorithmic component. The GUI component aims to provide a convenient environment for executing the algorithms by allowing users to enter input data and enabling users to view and analyze simulation results "on the fly". Users can see both the current and future network traffic states as generated by the state estimation and state prediction modules, respectively. Traffic statistics are provided at both the link and network levels. Also available are performance plots of the short-term and long-term consistency checking modules. Other features include the ability to view paths, temporal demand pattern, as well as attributes of nodes, links and the network.

2.5.2.1 Processed data information

The STCC, LTCC, and ODE in DYNASMART-X use observation data from different numbers of intervals. So the data is processed using two data interface procedures. One is External XML Data Interface, which obtains the detector data (count, speed, and occupancy) from the online XML website which specifies the

XML specific data schema every time interval, and then writes to an internal XML data file-DataInterface.xml (density), which keeps latest several durations data. The other is Internal XML Data Interface, which is used by STCC, LTCC, and ODE in DYNASMART-X to read the observed data based on their running time from internal XML data file-DataInterface.xml. The flowchart of the data processing procedure is shown in figure 2-4. It describes a XML data interface between DYNASMART-X system and external real world. Through the data interface, the real-time measurements during every observation interval are capable of being continuously provided to the simulation system.

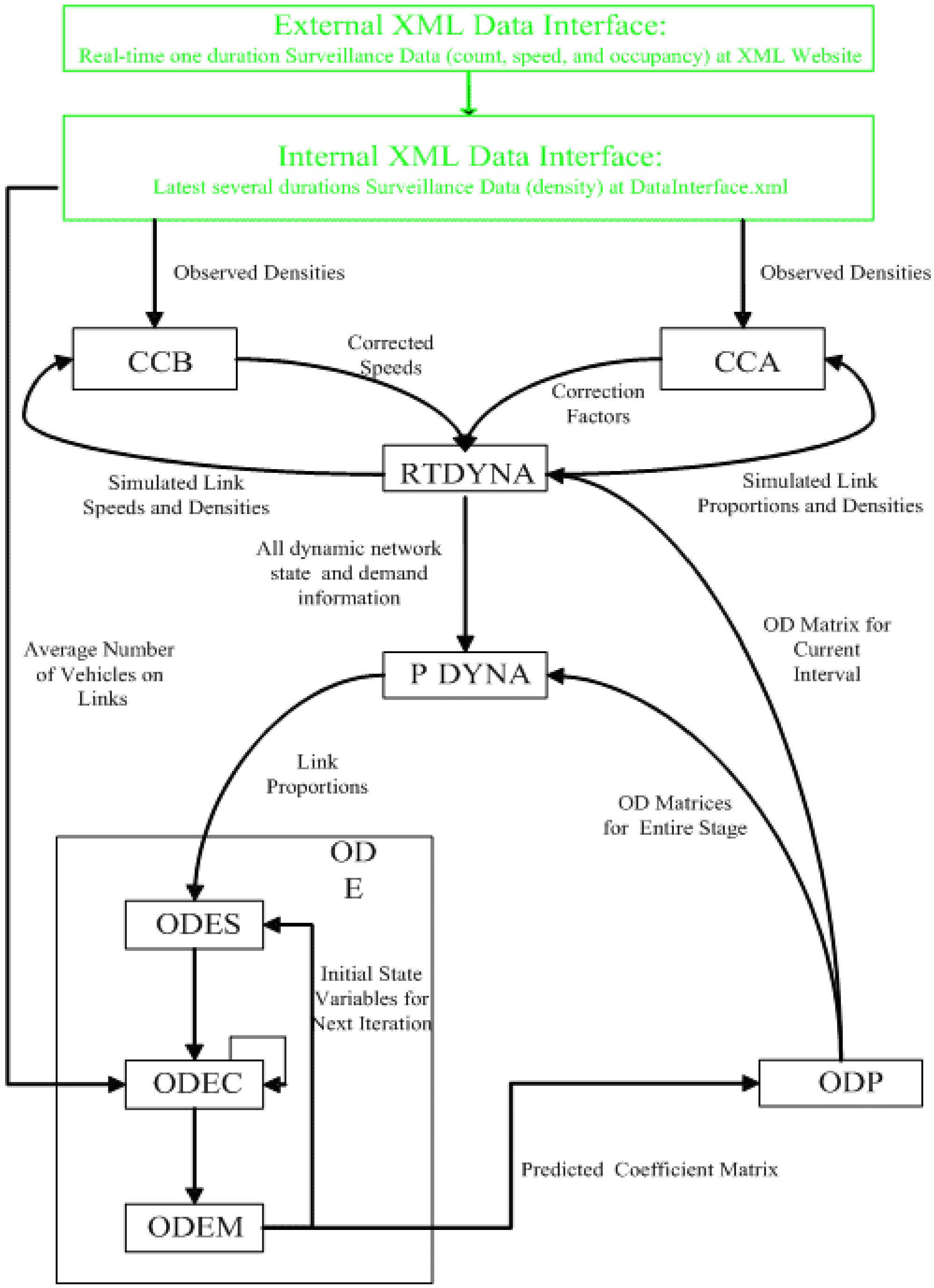


Figure 2-4 Surveillance Data Processing in DYNASMART-X

Figure 2-5 depicts the high-level structure of the system and the basic data flow model. The simulation proceeds in a rolling horizon fashion (Peeta and Mahmassani, 1995). The state estimation (RTDYNA) is executed periodically (every *assignment interval*), and continuously provides up-to-date estimates of the current state of the network. The state prediction (PDYNA) is executed less frequently (every *roll period*) and projects the current network state for a period in the future (the *stage length*), and the incorporated Multiple User Class (MUC) algorithm provides the route guidance information (Peeta and Mahmassani, 1995). The OD Estimation and Prediction modules provide the time-dependent OD desires in the network to be used in the simulation-assignment procedures of the state estimation and prediction. They also run periodically. The Consistency Checking modules interface with the surveillance data collected from sensors and probes in the network, and correct some of the state estimation variables for discrepancies between the estimated values and the measured ones. They run periodically, and their respective periods are design parameters that can vary according to the particular network being modeled and the experimental setting.

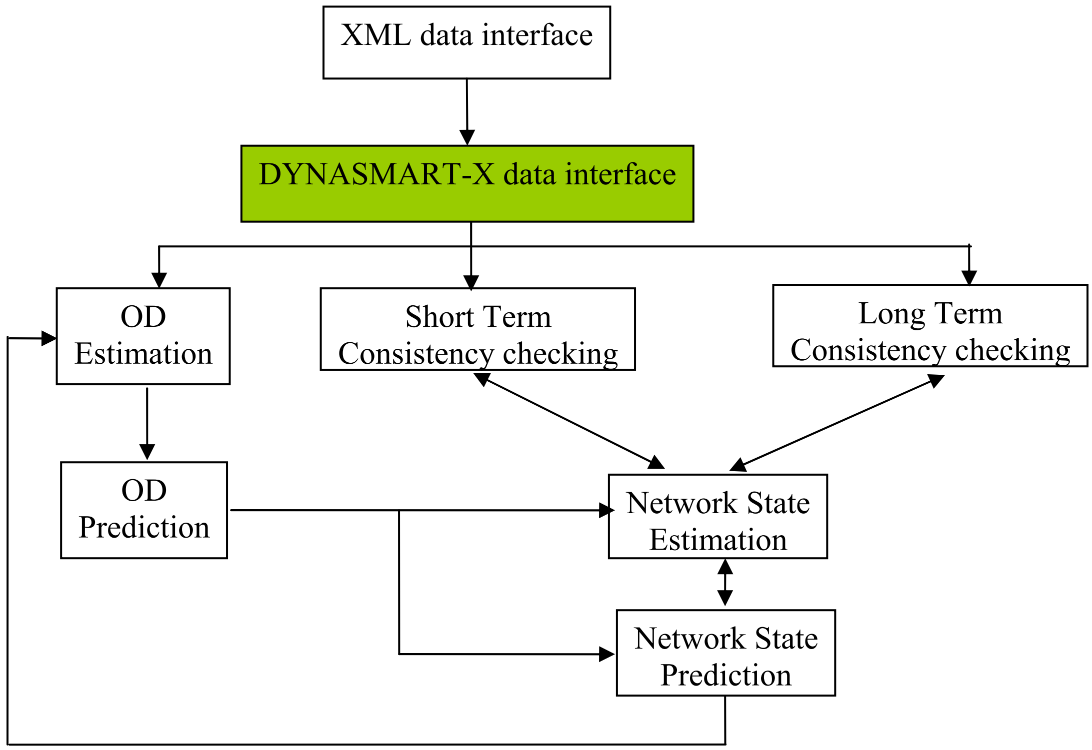


Figure 2-5 XML Real-Time Data Interface

2.5.2.2 Multiple scenario prediction methodology

DYNASMART-X includes both real-time traffic estimation and prediction elements. It interfaces with an external environment consisting of the entire traffic network in an urban area with all its static and dynamic elements, which include the network topology and geometry, traffic control devices, human users with their complex behavioral structure, in addition to the information being disseminated to users by various means. The information element is of central importance in defining the operational role of DYNASMART-X as a predictive, rather than merely reactive real-time system, since it also contributes to the information being supplied to users and traffic control systems (Mahmassani, 1998). There are two instances in DYNASMART-X where message-based asynchronous communication is useful between DYNASMART- X and the external environment. The first is the publisher/subscriber communication pattern that links DYNASMART-X to its clients (or to external systems). The second is event notification in the reverse direction, i.e. to notifying the simulation engine of external events that need to be processed, for example, incidents and VMS status changes. Figure 2-6 shows the different message channels that are implemented. First, the subscribed client (the GUI in the figure, e.g. the TMC operator) is notified regularly about internal events that are occurring in the engine. For example, that the current PDYNA instance (state prediction instance) has finished execution. Once notified, the GUI can take the appropriate action. For example, it can contact data-broker (DBK) to get the latest estimate of the network state, and update the display.

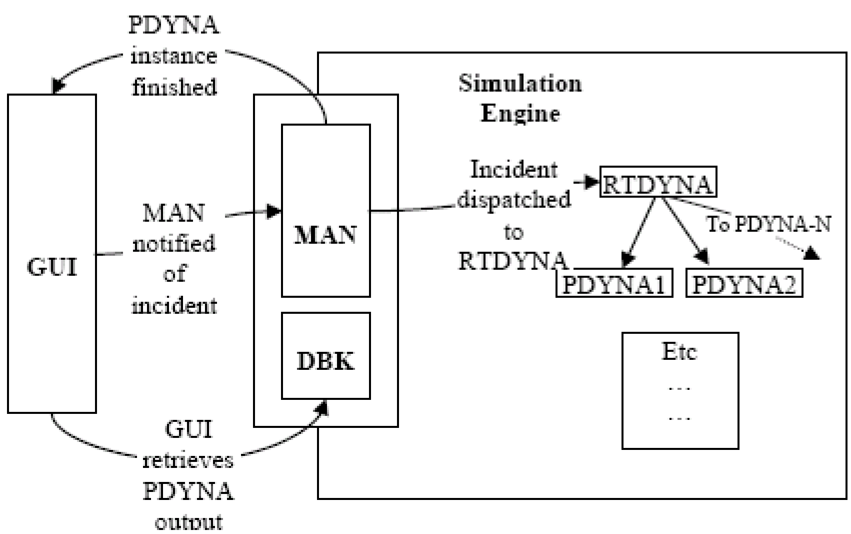


Figure 2-6 Message Channels between Engine and Clients

On the other hand, once an incident is detected on the traffic network, the user inputs the incident parameters via the GUI, and a notification message is sent immediately to the management component (MAN), the "control center", describing this incident. MAN immediately dispatches the incident information to RTDYNA (also via a message, thus RTDYNA is implemented as a Message Target), which places this information into a queue. Before the next instance of RTDYNA is launched, this queue is scanned and the traffic events are processed as necessary. The state estimation (RTDYNA) is executed periodically (every *assignment interval*), and continuously provides up-to-date estimates of the current state of the network. RTDYNA then transfers the entire set of state variables that define the network and the traffic conditions at that instant to PDYNA, which will use the current network state as a starting point to project a period in the future (the *stage length*) (Mahfoud 2005).

Multiple PDYNA instances allow the operator to evaluate multiple traffic control/management strategies in real-time fashion. From the standpoint of evaluating control strategies online, multiple instances of PDYNA can be initiated and executed in parallel, with each taking the same initial network state but different control and information provision strategy (Figure 2-7). There are two modes in which multiple PDYNA can be activated; sequential mode and real-time mode. Sequential mode runs multiple instances of PDYNA sequentially, while the real-time mode will simultaneously run the multiple instances of PDYNA under the real-time clock. The sequential mode implements the rolling horizon logic in an artificial way that preserves the logical dependence between the modules, but without enforcing any timing constraints on their execution. It is intended for off-line testing of the system. The real-time mode implements the rolling horizon simulation logic with all the applicable timing constraints. Figure 2-7 shows the different running time for sequential mode and real-time mode. The running time of sequential mode is the summation of the PDYNA0(t_1) and PDYNA1(t_2) while the running time of real-time mode depends on the maximum running time of PDYNA0(t_1) and PDYNA1(t_2).

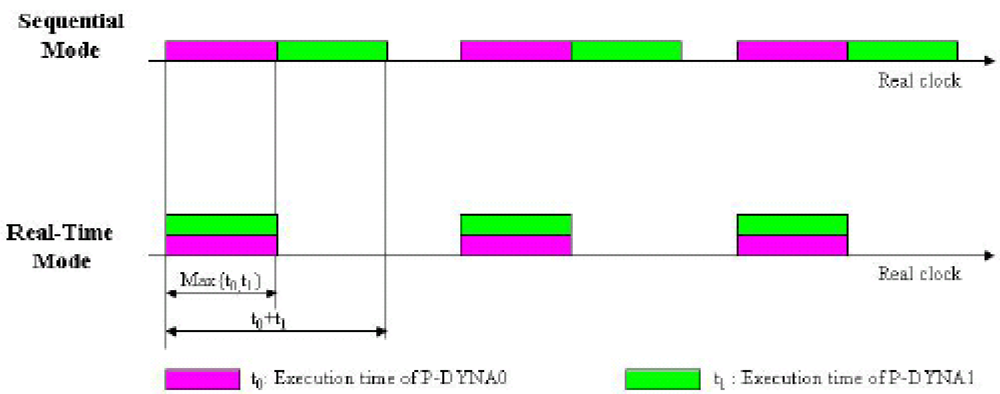


Figure 2-7 Comparison of Execution Time of Sequential Mode & Real-Time Mode

2.5.2.3 Real-time traffic management decision support methodology

The ability to evaluate multiple traffic management strategies in quasi real-time using DYNASMART-X can aid in decision support and can improve the ability of the traffic management center to respond to unfolding situations including incidents, congestion and other unexpected events, through provision of traffic information to travelers and deployment of various control measures. To summarize the process from a TMC operator's standpoint, the DYNASMART-X simulator runs as it normally would, making predictions and estimations on the basis of real-time information. When a change in traffic conditions occurs, the simulator will change to reflect the conditions on the basis of the real information it is receiving. If an unplanned disturbance occurs, the operator at the TMC can inform DYNASMART by making changes to reflect the disturbance (e.g. implementing an incident of corresponding severity). This allows the simulator to adjust for changes to the physical network or control processes and more accurately replicate field conditions. The ability to assess multiple alternative management strategies is desired when traffic conditions worsen or an unplanned event occurs. When this is the case, the TMC operator can construct strategies or plans for mitigating the traffic problems. For example, if an accident occurs and two lanes on a 3-lane highway are closed; the TMC operator informs DYNASMART that only one lane is functioning at the incident location and can then develop strategies for routing vehicles around the accident or altering the control plans around the incident location. Once the TMC operator has devised response strategies, they can be implemented as different instances of PDYNA. Each instance of PDYNA receives information from RTDYNA as described previously. The results each PDYNA instances provide the TMC operator the ability to see the results of implementing each of the alternate strategies. Once the TMC operator has selected a strategy and implemented it in the field, the TMC operator inputs the changes in the DYNASMART GUI to reflect the changes that were made in the field.

2.6 Summary

An accurate OD matrix plays a critical role in applications of DTA models to support advanced transportation management and traveler information systems. Because an OD matrix is prohibitively expensive to obtain directly, it is often estimated using measurement data from the traffic network. Because each observation link may contain different information, the proper deployment of the sensors as well as the use of statistically based OD estimators are essential for successful traffic management system. Clearly, uncertainty is associated with the demand estimation or prediction. It is important to take the uncertainty into the sensor location model formulation in order to exploit a set of robust sensor locations in terms of providing high quality of estimated demand with regard to extreme cases. In this chapter, the relevant background concerning OD demand estimation/prediction methods under static/dynamic traffic assignment are reviewed, followed by an overview of different existing sensor location approaches, and overview of stochastic programming approaches with the incident generation methods. Then, the simulation based dynamic traffic assignment tool, DYNASMART (offline version and online version) is introduced. Finally, a multiple scenario prediction methodology of DYNASMART-X is presented.

Chapter 3 Finding Near-Optimal Sensor Locations for Large-Scale Network Under Deterministic Network Condition

3.1 Introduction

The sensor location problem could be viewed from a value of information perspective. Sensors continuously provide information that help characterize the status of the network. Using this information in conjunction with "knowledge" (i.e. historical data, previous estimation or prediction outputs) could enhance a model's estimation and prediction performance (see Figure 3-1). Adding sensors to the network at specific locations could be evaluated with regard to the additional value that these sensors provide to the ability to estimate and predict network flow patterns (e.g. OD demands, path flows, link flows, point speeds), provide travel time information, or provide better control strategies.

Ideally, one would want sensors on all the links in the network. This would reduce the error associated with the state estimation to the system error. Focusing on the sensor location problem, the principal goal of this chapter is to identify the locations which provide the most value given a limiting constraint on the number of sensors, and propose an associated mathematical model and efficient solution procedure based on dynamic traffic assignment methodology to strategically deploy the given number sensors in large scale road networks. The solution procedure operates in two steps. In step one, a restricted candidate list (RCL) is generated from choosing a set of top candidate locations sorted by link flows. A predetermined number of links is randomly selected from the RCL according to a link-independent rule. In step two, the selected candidate locations generated from step one are evaluated in terms of the magnitude of the flow variation reduction and O-D flow coverage using archived historical and simulated traffic data.

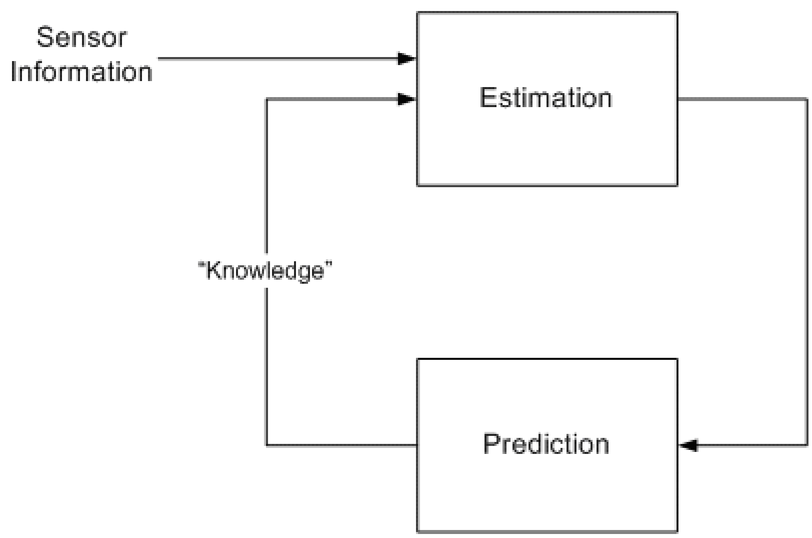


Figure 3-1 Estimation and Prediction Enhancement Information

Two sensor location problems are proposed and analyzed in this chapter. One is to find a minimal number of sensors and locations to cover a percentage (0%-100%) of the time-dependent traffic flows on the network. The other is to identify a set of given number sensor locations that maximize the coverage of origin-destination (O-D) flows of the road network, while minimizing the uncertainty of the estimated time-dependent O-D demand matrix. Considering demand coverage and uncertainty reduction simultaneously, the second case is formulated as a bi-objective problem.

The rest of this chapter comprises four sections. Section 3.2 presents a framework for approaching the sensor location problem and discusses models that can be used for both of the cases with and without budgetary constraint. Section 3.3

includes an analysis that illustrates the information gains and trade-offs associated with various sensor location schemes. Section 3.4 examines the results produced by the proposed models. Section 3.5 summarizes the chapter.

3.2 Conceptual Framework

The sensor location problem is more than a simple coverage problem. Even if every link of the entire network has a point sensor installed, the network path flows may still not be uniquely determined. In the case without budgetary constraint, the objective of the proposed problem is to minimize the number of sensors while covering all the O-D flows during each observation time interval in the network. This case can be categorized as a set covering optimization problem. However, in most cases, it might be difficult to have the sensors fully installed on the entire road network due to the budgetary constraint. With the given number of sensors, the goal becomes to capture the network traffic flows as much as possible and minimize the network performance uncertainty using the information brought by every sensor. Thus the importance of a location depends on the value of information/knowledge that it can bring to the problem.

3.2.1 General Least Squares OD Demand Estimator

Consider a network with n_{obs} observation link, n_{Zone} zones ($i, j \in n_{Zone}$) and n_{OD} OD pairs. τ represents the departure interval of each OD pair. Assume the true OD demand $d^{\tau}_{(i,j)}$ can be decomposed into three components, a *priori* estimation $\hat{d}^{\tau}_{(i,j)}$, structural deviation $\chi^{\tau}_{(i,j)}$ between the actual demand and *a priori* estimated demand of OD pair (i, j) during departure time interval τ and random error $\xi^{\tau}_{(i,j)}$ where $\xi^{\tau}_{(i,j)} \sim N(0, P_{\hat{D}}(-))$ is the estimation error (Zhou 2004). The linear combination can be described as follows:

$$d^{\tau}_{(i,j)} = \hat{d}^{\tau}_{(i,j)} + \chi^{\tau}_{(i,j)} + \xi^{\tau}_{(i,j)} \tag{3.1}$$

Structural deviation $\chi^{\tau}_{(i,j)}$ can capture different factors such as special events, weather conditions, incidents, and temporary physical changes of the transportation network, etc.

Assume the relationship between the unknown OD flow and measurements can be expressed as a combination with a random, additive measurement error ε. The measurement process is modeled as:

$$c_{l,t} = \sum_{i,j} \sum_{p=t-q}^{t} ((\hat{L}P_{(l,t),(i,j,p)} + \zeta_{(l,t),(i,j,p)}) \cdot d^{p}_{(i,j)}) + \varepsilon^{t}_{l} \tag{3.2}$$

Where $\hat{L}P_{(l,t),(i,j,p)}$ is the link proportions, $\zeta_{(l,t),(i,j,p)} \sim N(0, \delta)$ is the assignment error The link flow may be composed of OD flows from different previous time interval including current time interval. The time lag q is determined by the length of time interval and maximal magnitude of travel time between an origin and a destination in the network.

Substituting (3.1) into (3.2),

$$c_{l,t} = \sum_{i,j} \sum_{p=t-q}^{t} ((\hat{L}P_{(l,t),(i,j,p)} + \zeta_{(l,t),(i,j,p)}) \cdot (\hat{d}^{p}_{(i,j)} + \chi^{p}_{(i,j)} + \xi^{p}_{(i,j)})) + \varepsilon^{t}_{l}$$

$$= \sum_{i,j} \sum_{p=t-q}^{t} ((\hat{L}P_{(l,t),(i,j,p)} \cdot \hat{d}^{p}_{(i,j)} + \hat{L}P_{(l,t),(i,j,p)} \cdot \chi^{p}_{(i,j)} + \zeta_{(l,t),(i,j,p)} \cdot d^{p}_{(i,j)} + \hat{L}P_{(l,t),(i,j,p)} \cdot \xi^{p}_{(i,j)})) + \varepsilon^{t}_{l}$$

$$= \sum_{i,j} \sum_{p=t-q}^{t} \hat{L}P_{(l,t),(i,j,p)} \cdot \hat{d}^{p}_{(i,j)} + \sum_{i,j} \sum_{p=t-q}^{t} \hat{L}P_{(l,t),(i,j,p)} \cdot \chi^{p}_{(i,j)} + \sum_{i,j} \sum_{p=t-q}^{t} (\zeta_{(l,t),(i,j,p)} \cdot d^{p}_{(i,j)} + \hat{L}P_{(l,t),(i,j,p)} \cdot \xi^{p}_{(i,j)}) + \varepsilon^{t}_{l}$$

Let $c^{'}_{l,t} = c_{l,t} - \sum_{i,j} \sum_{p=t-q}^{t} \hat{L}P_{(l,t),(i,j,p)} \cdot \hat{d}^{p}_{(i,j)}$, $\upsilon_{l,t}$ denotes the combined error; it has

$$c^{'}_{l,t} = \sum_{i,j} \sum_{p=t-q}^{t} \hat{L}P_{(l,t),(i,j,p)} \cdot \chi^{p}_{(i,j)} + \upsilon_{l,t} \tag{3.3}$$

For convenience, stage symbol t is dropped off here for the model derivation.

In the matrix form, Eq.(3.3) can be written as

$$\mathbf{C} = \hat{\mathbf{H}} \cdot \mathbf{D} + \boldsymbol{\varepsilon} \tag{3.4}$$

Where $\mathbf{C}\,(n_{Obs} x 1)$ is an observation vector, $\hat{\mathbf{H}}\,(q\ x\ n_{Obs} x\ n_{OD})$ is an assignment matrix that mapping demand $\mathbf{D}$ into link counts $\mathbf{C}$. $\mathbf{D}$ $(q\ x\ n_{OD}\ x\ 1)$ is a structural deviation vector, $E(\boldsymbol{\varepsilon}) = 0$, $E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{T}) = \mathbf{R}$, $\mathbf{R}$ is a known symmetric, positive-definite matrix.

Eq.3.3 represents a general non-linear relation between the deviation of link traffic counts and the unknown demand structural deviation including the confounded error terms. Because of the non-linearity, the combined error $\upsilon_{l,t}$ is not white noise. For the reason of the focus of this research is sensor locations, we assume that the sum of error terms and the interaction terms is a normal distribution with zero mean and unknown dispersion, but since the interactions and the error terms are ignored, it may result in possible inaccuracy results and an biased and inefficient estimator (an efficient estimator by definition is the one with the lowest variance among all unbiased estimators, and it will be further discussed later in this chapter). Note that Eq.3.3 uses the structural deviation as the state variable in order to capture trip patterns and their temporal and spatial variations. Moreover, under the normal distribution of the traffic variables, i.e. link counts, OD flows, etc, the deviation formation is more amenable to approximate the normal distribution than the traffic variables because they can take both positive and negative values (Askok et al.. 2000).

The objective is to minimize the sum of squared residuals $\boldsymbol{\varepsilon}$. Its use in the context of GLS estimation does not require any distributional assumptions and according to the *general-least-square* estimation (the notations referred to section 3.3),

$$J = \arg\min(\mathbf{C} - \mathbf{H}\hat{\mathbf{D}}(-))^T \mathbf{R}^{-1} (\mathbf{C} - \mathbf{H}\hat{\mathbf{D}}(-)) \tag{3.5}$$

By setting $\frac{\partial J}{\partial \hat{D}(-)} = 0$, the resultant closed form GLS estimator is

$$\hat{\mathbf{D}}(-) = (\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1} \mathbf{C} \tag{3.6}$$

Assuming the measurement errors are uncorrelated, e.g. $\mathbf{R} = \mathbf{I}$, it is easy to prove that

$$\hat{\mathbf{D}}(-) = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{C} \tag{3.7}$$

Note that for any matrix $\mathbf{H}$, the $rank(\mathbf{H}) = rank(\mathbf{H}^T\mathbf{H}) = rank(\mathbf{H}\mathbf{H}^T)$, such that if matrix $\mathbf{H}$ is of full rank, then the least squares solution $\hat{\mathbf{D}}(-)$ is unique and minimizes the sum of squared residuals. In another word, the link counts on each observed link needs to be linearly independent with each other.

According to Aitken's theorem (1935), the GLS estimator $\hat{\mathbf{D}}(-)$ is the minimum variance linear unbiased estimator in the generalized regression model. Cascetta (1984) discussed the statistical properties of the GLS estimator (referred to as Aitken estimator) and analyzed two cases, stochastic and deterministic observation flow with simulation data in terms of mean square error (MSE/Risk) and generalized mean square error (GMSE). He pointed out that if the *a priori* estimator and assignment model were correctly specified, the Aitken estimator with inactive inequality constraint is the best linear unbiased estimator (BLUE). In this research, the simulated assignment matrices are assumed to represent the actual one.

Using time-varying weighting matrices $\mathbf{K}$ and $\mathbf{K}^{'}$, the recursive form can be expressed as

$$\hat{\mathbf{D}}(+) = \mathbf{K}^{'}\hat{\mathbf{D}}(-) + \mathbf{KC} \tag{3.8}$$

Since

$$\begin{aligned} \hat{\mathbf{D}}(+) &= \mathbf{D} + \widetilde{\mathbf{D}}(+) \\ \hat{\mathbf{D}}(-) &= \mathbf{D} + \widetilde{\mathbf{D}}(-) \end{aligned} \tag{3.9}$$

Substituting (3.4) and (3.8) into (3.9), it gets

$$\begin{aligned} \widetilde{\mathbf{D}}(+) &= \mathbf{K}^{'}(\mathbf{D} + \widetilde{\mathbf{D}}(-)) + \mathbf{K}(\mathbf{HD} + \boldsymbol{\varepsilon}) - \mathbf{D} \\ &= (\mathbf{K}^{'} + \mathbf{KH} - \mathbf{I})\mathbf{D} + \mathbf{K}^{'}\widetilde{\mathbf{D}}(-) + \mathbf{K}\boldsymbol{\varepsilon} \end{aligned} \tag{3.10}$$

$\hat{\mathbf{D}}(-)$ or $\hat{\mathbf{D}}(+)$ is unbiased. That is

$$\left.\begin{aligned} \mathbf{E}(\hat{\mathbf{D}}(+)) = \mathbf{E}(\mathbf{D} + \widetilde{\mathbf{D}}(+)) = \mathbf{D} + \mathbf{E}(\widetilde{\mathbf{D}}(+)) \\ \mathbf{E}(\hat{\mathbf{D}}(-)) = \mathbf{E}(\mathbf{D} + \widetilde{\mathbf{D}}(-)) = \mathbf{D} + \mathbf{E}(\widetilde{\mathbf{D}}(-)) \end{aligned}\right\} \Rightarrow \begin{cases} \mathbf{E}(\widetilde{\mathbf{D}}(+)) = \mathbf{0} \\ \mathbf{E}(\widetilde{\mathbf{D}}(-)) = \mathbf{0} \end{cases} \tag{3.11}$$

By definition, $E(\boldsymbol{\varepsilon}) = 0$, (3.10) and (3.11) give

$$\mathbf{K}^{'} = (\mathbf{I} - \mathbf{KH}) \tag{3.12}$$

Substituting (3.12) into (3.9)

$$\widetilde{\mathbf{D}}(+) = (\mathbf{I} - \mathbf{KH})\widetilde{\mathbf{D}}(-) + \mathbf{K}\boldsymbol{\varepsilon} \tag{3.13}$$

By definition, the *posteriori* error variance covariance matrix($n_{OD}\ x\ n_{OD}$)

$$\mathbf{P}_{\hat{\mathbf{D}}}(+) = \mathbf{E}(\widetilde{\mathbf{D}}(+) - \mathbf{E}(\widetilde{\mathbf{D}}(+))(\widetilde{\mathbf{D}}(+) - \mathbf{E}(\widetilde{\mathbf{D}}(+))^{\mathbf{T}} \\ = \mathbf{E}(\widetilde{\mathbf{D}}(+)\widetilde{\mathbf{D}}(+)^{\mathbf{T}}) \tag{3.14}$$

Substituting (3.13) into (3.14),

$$\mathbf{P}_{\hat{\mathbf{D}}}(+) = ((\mathbf{I} - \mathbf{KH})\widetilde{\mathbf{D}}(-) + \mathbf{K}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon})((-\mathbf{KH})\widetilde{\mathbf{D}}(-) + \mathbf{K}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathbf{T}} \\ = (\mathbf{I} - \mathbf{KH})\mathbf{P}_{\hat{\mathbf{D}}}(-)(\mathbf{I} - \mathbf{KH})^{\mathbf{T}} + \mathbf{KRK}^{\mathbf{T}} \tag{3.15}$$

To minimize $\mathbf{P}_{\hat{\mathbf{D}}}(+)$, the first-order optimization condition (FOC) needs to be satisfied,

$$\frac{\partial \mathbf{P}_{\hat{\mathbf{D}}}(+)}{\partial \mathbf{K}} = -\mathbf{2}(\mathbf{I} - \mathbf{KH})\mathbf{P}_{\hat{\mathbf{D}}}(-)\mathbf{H}^{\mathbf{T}} + \mathbf{2KR} = \mathbf{0} \tag{3.16}$$

Thus, the optimal weight matrix, which is referred to as Kalman gain matrix is

$$\mathbf{K} = \mathbf{P}_{\hat{\mathbf{D}}}(-)\mathbf{H}^{\mathbf{T}}(\mathbf{HP}_{\hat{\mathbf{D}}}(-)\mathbf{H}^{\mathbf{T}} + \mathbf{R})^{-1} \tag{3.17}$$

As an incremental algorithm, Kalman filter algorithm is used to solve a least square problem in a real-time context, $\mathbf{K}$ is a Kalman gain matrix ($n_{OD}\ x\ n_{obs}$). Substituting (3.17) into (3.15), the minimal updated variance covariance matrix is

$$\mathbf{P}_{\hat{\mathbf{D}}}(+) = (\mathbf{I} - \mathbf{KH})\mathbf{P}_{\hat{\mathbf{D}}}(-) \tag{3.18}$$

A simple form of Kalman gain matrix can be expressed as

$$\mathbf{K} = \mathbf{P}_{\hat{\mathbf{D}}}(+)\mathbf{H}^{\mathbf{T}}\mathbf{R}^{-1} \tag{3.19}$$

Equation (3.15) can be also expressed as

$$\mathbf{P}_{\hat{\mathbf{D}}}^{-1}(+) = \mathbf{P}_{\hat{\mathbf{D}}}^{-1}(-) + \mathbf{H}^{\mathbf{T}}\mathbf{R}^{-1}\mathbf{H} \tag{3.20}$$

More detailed derivations and analysis about the optimal estimation and filtering relationship can be found in Gelb (1974).

If we assume that the measurement error is independent, then **R** is a diagonal matrix. So, Equation (3.19) can be written as

$$\mathbf{K} = \frac{\mathbf{P}_{\hat{\mathbf{D}}}(+)\mathbf{H}^{\mathrm{T}}}{\mathbf{R}} \tag{3.21}$$

The matrix $\mathbf{H}$ is a mapping matrix, mapping the OD demand flow to the link counts; if it is assumed to be an identity matrix, one would get

$$\mathbf{K} = \frac{\mathbf{P}_{\hat{\mathbf{D}}}(+)}{\mathbf{R}} \tag{3.22}$$

From (3.4).(3.8), (3.9) and (3.12), it can get

$$\widehat{\mathbf{D}}(+) = \widehat{\mathbf{D}}(-) + \mathbf{K}(\mathbf{C} - \mathbf{H}\widehat{\mathbf{D}}(-)) \tag{3.23}$$

The Kalman filter method is a recursive approach for estimating an unknown state vector $\mathbf{D}_k$ at each instance k, $k = 1,2,.....$ in a discrete linear stochastic system that gives a linear, unbiased and minimum error variance estimate. Eqs. (3.17), (3.18) and (3.23) can be derived from the standard Kalman filtering procedure. Detailed derivations and analysis about the optimal estimation and filtering relationship can be found in Gelb (1974). Thus, the sensor location problem becomes a traffic state learning process (Figure 3-2, similar to the sequential algorithm of Chui & Chen (1991)) that seeks to locate sensors which recursively add valuable information to update estimates (in terms of mean and variance) on the network traffic states. The key question is how to characterize the value of additional information from a new detector in traffic state estimation and prediction.

3.2.2 Link Kalman Gain and Uncertainty

Link Kalman gain K_l in the sensor location problem can be interpreted as the summation of information gain brought by each O-D flow that intercepted by link *l*. A simple form of Kalman gain matrix can be expressed as Eq.(3.21).

$$\mathbf{K} = \frac{\mathbf{P}_{\hat{\mathbf{D}}}(+)\mathbf{H}^{\mathbf{T}}}{\mathbf{R}}$$

Apparently, it is "proportional" to the network estimate uncertainty and "inversely proportional" to the measurement noise. Thus the goal of locating sensors would then be to identify those places with less measurement noises and additional measurements that minimize uncertainty resulting covariance matrix $\mathbf{P}_{\hat{\mathbf{D}}}(+)$. Eq.(3.18) says that given *a priori* demand uncertainty, large link information gain provides large uncertainty reduction. The above covariance updating formula clearly links the *a priori* uncertainty and a posteriori uncertainty, and $\mathbf{KH}$ measures the degree of uncertainty reduction due to inclusion of new measurements. New measurements can come from a single sensor or multiple sensors. $\mathbf{KH}$ can be viewed as a matrix specifying the value of additional information.

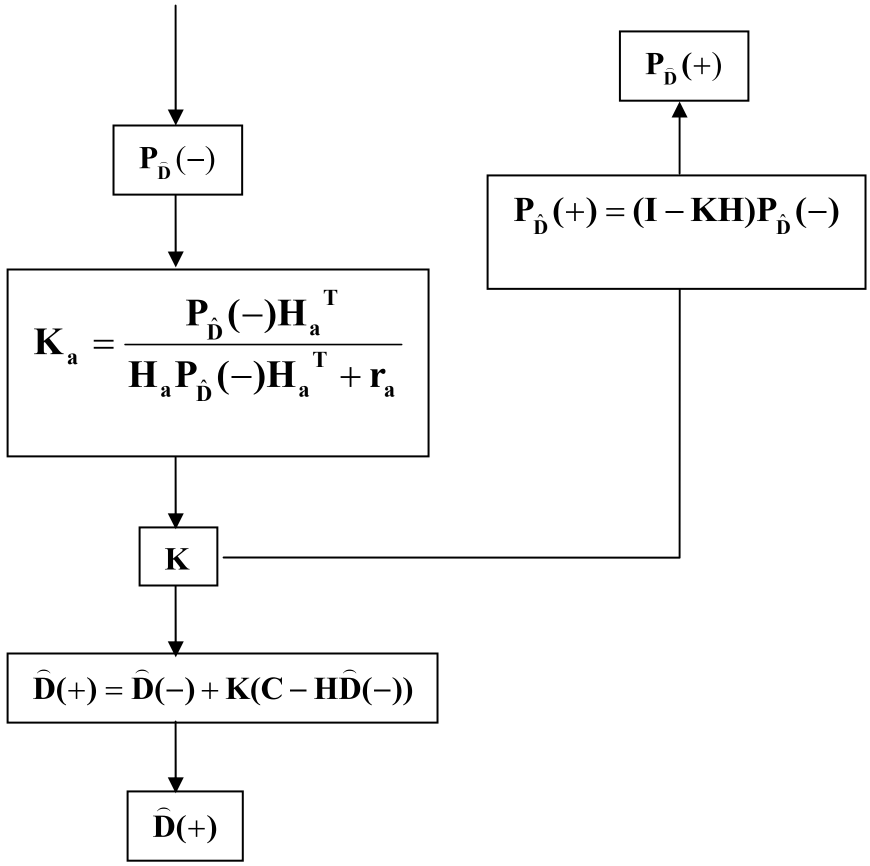


Figure 3-2 State Learning Process in Sensor Location Problem

Eq.(3.17) shows that the gain matrix is more sensitive to the measurement error than the demand uncertainty for each unknown O-D flows and covariance between different unknown flows. Moreover it discloses that the product of $\mathbf{P}_{\hat{\mathbf{D}}}(-)\mathbf{H}^T$ is more likely to be large and more uncertainty reductions are obtainable if selected locations can intercept those O-D pairs with large variances, or more O-D pairs. Another issue about the weighting matrix $\mathbf{K}$ is the inverse of ($\mathbf{H}\mathbf{P}_{\hat{\mathbf{D}}}(-)\mathbf{H}^T + \mathbf{R}$). If one only considers $\mathbf{H}\mathbf{H}^T$ for multiple possible sensors, the inverse of $\mathbf{H}\mathbf{H}^T$ specifies the correlation of measurements among multiple sensors.

$\mathbf{HP}_{\hat{\mathbf{D}}}(-)\mathbf{H}^T$, furthermore, describes the measurement correlation on the basis of existing estimate variance and covariance. If $\mathbf{HP}_{\hat{\mathbf{D}}}(-)\mathbf{H}^T$ is large, it means that either new senor data could be highly correlated with each other or they are correlated with the current estimate, then the inverse is small and the weight factor becomes insignificant. Clearly, the more measurement error the less uncertainty reduction there will be for the estimates.

If the link counts are statistically assumed independent, the inversion of matrix $\mathbf{HP}_{\hat{\mathbf{D}}}(-)\mathbf{H}^T + \mathbf{R}$ can be avoided (Chui & Chen (1991)). For each individual link, $(\mathbf{H}_l\mathbf{P}_{\hat{\mathbf{D}}}(-)\mathbf{H}_l^T + \mathbf{R}_l)^{-1}$ may be obtained by a scalar inversion. Otherwise, the multivariate minimization procedures, such as the Davidon-Flecher-Powell procedure (Scales 1985) that used to solve unconstrained GLS problem without through matrix inversion can be used to yield $(\mathbf{HP}_{\hat{\mathbf{D}}}(-)\mathbf{H}^T + \mathbf{R})^{-1}$.

3.2.3 The Assignment Matrix

The assignment matrix maps the O-D flows onto the link counts. The two main classes of assignment process are proportional assignment that the assignment matrices are independent with O-D flow and equilibrium assignment that link flows depend on the link capacity. Clearly, it is an import input to the sensor location problem. In the context of dynamic traffic assignment (DTA), the assignment matrix is not constant, and themselves are dependent with the unknown time-dependent demand flows. It captures three aspects of a traffic network: the network topology, the route choice model and the travel time across the network (Bierlaire and Crittin 2004).

Let $\alpha^h_{l,w,t}$ represents the fraction of w^{th} O-D flow that left its origin at departure time t and traversed over link l during observation interval h. $\alpha^h_{l,p,t}$ defines time-dependent link path indicator. It equals 1 if path flow p left origin at departure time t and traversed over link l during observation interval. $q_{p,t}$ denotes path flow choice probability that select path p during departure time t. Cascetta *et. al* (1993) shows the relationship between link flow proportion $\alpha^h_{l,w,t}$ and link path incidence as follows:.

$$\alpha^h_{l,w,t} = \sum_{p \in K} (\alpha^h_{l,p,t} * q_{p,t}) \tag{3.24}$$

Eq.(3.24) shows that the assignment matrix is determined by the route choice fraction and network traffic flow propagation. Based on the assumptions that the vehicles are uniform distribution in a packet and travel times are observable, Cascetta et. al (1993) derived a relationship between the link path incidence and travel time. However, it has different error sources that may lead network representation deviating from the actual network causing erroneous travel time estimation and/or incorrect path flow choice split. Those include (1) demand estimation errors (2) path estimation errors (3) traffic propagation errors (4) internal traffic model structure errors (5) on-line data observation errors (Doan *et al* 1998). Those errors may result in biased and inconsistency O-D estimations. Ashok *et. al* (2002) analyzed conditions that part or all of the travel times are endogenous. They proposed two approaches to model the stochasticity of the assignment matrix. In the first approach, a random error v^h_t is introduced to the unknown actual assignment matrix $\alpha^{h}_{l,w,t}$, such that $\alpha^{h}_{t} = \hat{\alpha}^{h}_{t} + v^{h}_{t}$. As an alternative approach induced from Eq.(3.24), the assignment matrix is defined as a function of travel time and route choice fraction, $\alpha^{h}_{t} = F(T_h, q_p)$.

Due to the computation complexity and intensiveness of the assignment matrix in a large scale network, the time-dependent assignment matrix is obtained from a dynamic traffic assignment model based simulation software DYNASMART-P (Mahmassani et. al 2000) in this research. The user-equilibrium (UE) and system-optimal(SO) procedures are integral components of DYNASMART-P (Peeta & Mahmassani 1995). The drivers in the network were assumed to take the paths consistency with those generated from the dynamic user equilibrium assignment.

3.2.4 Gain Collection and O-D Demand Coverage

The sensor location problem is a network design problem while the traffic pattern and behavior are dynamic that could be influenced by different factors, such as land use, special events, weather, etc. It is a trade-off to the decision makers to make his/her decisions between the system uncertainty reduction and O-D flow coverage. An O-D pair w is regarded as being *covered* if path flow $f^{t}_{w,p}$ of that O-D pair is intercepted by at least one of sensors in the network, where $f^{t}_{w,p}$ is defined as a path flow of O-D pair w along path p departed from origin during time interval t. If an O-D pair is uncovered, the demand of that particular O-D pair is not impacted by the network sensors and thus cannot be inferred from the observed flow.

The sensor location problem in this chapter is considered as a bi-objective problem under the assumption of recurrent traffic condition. One of the objectives is to minimize the demand uncertainty and the other is to maximize the O-D demand coverage as follows:

$$\begin{cases} Min\ F_1(Z) = Min\left\{ \sqrt{\dfrac{\sum_{\tau}\sum_{w}\left(d_w^{\tau} - \hat{d}_w^{\tau}(+)\right)^2}{|W| \times T}} \right\} & (3.25) \\ Max\ F_2(Z) = Max \sum_{\tau \le t \in T} \sum_{l \in M_0} \sum_{w \in W} \hat{d}_w^{\tau}(-) * I(h_{l,w}^{\tau,t}) * z_l & (3.26) \end{cases}$$

Eq.(3.25) is minimizing the deviation between the actual O-D demand and the estimated *a posteriori* demand across over all of the O-D pairs and the whole planning time horizon. Eq.(3.26) is maximizing the O-D flow coverage by the network sensors. Note that $I(\bullet)$ is an indicator function that assures an O-D demand flow departed from an origin during time interval τ is counted only once in time t. However, the ground truth O-D trips usually unknown. Eq.(3.25) thus can be translated into maximizing total link information gains as discussed earlier as $Max\ F_1(Z) = \sum_{\tau \in T} \sum_{l \in A} \sum_{w \in W} (k_{l,w}^{\tau} * z_l)$.

Figure 3-3 conceptually shows that the efficient frontier of domain $\mathbf{R^L}$, which is the non-dominated solution set, yields set of possible location sets depending on the preference of the decision maker to the link information gains or OD flow coverage. Various methods, such as weighting objectives method, hierarchical optimization method, trade-off method, global criterion method, goal programming method, min-max optimum, method of distance functions, have been developed to find the Pareto optimal set. The linear weighting method exploring the efficient frontier is used in this study (conceptually showed in figure 3-2), which helps the decision maker determine the different weight combinations to get the best compromise solution set Z^*. Specifically if the decision maker is preferring the O-D flow coverage in terms of the dynamic traffic information and control operation to reducing the system uncertainty, the ratio of $\frac{w_2}{w_1} > 1$. Otherwise if the decision maker is more concerned about minimizing the system uncertainty based on the a *priori* demand, the ratio of $\frac{w_2}{w_1} < 1$.

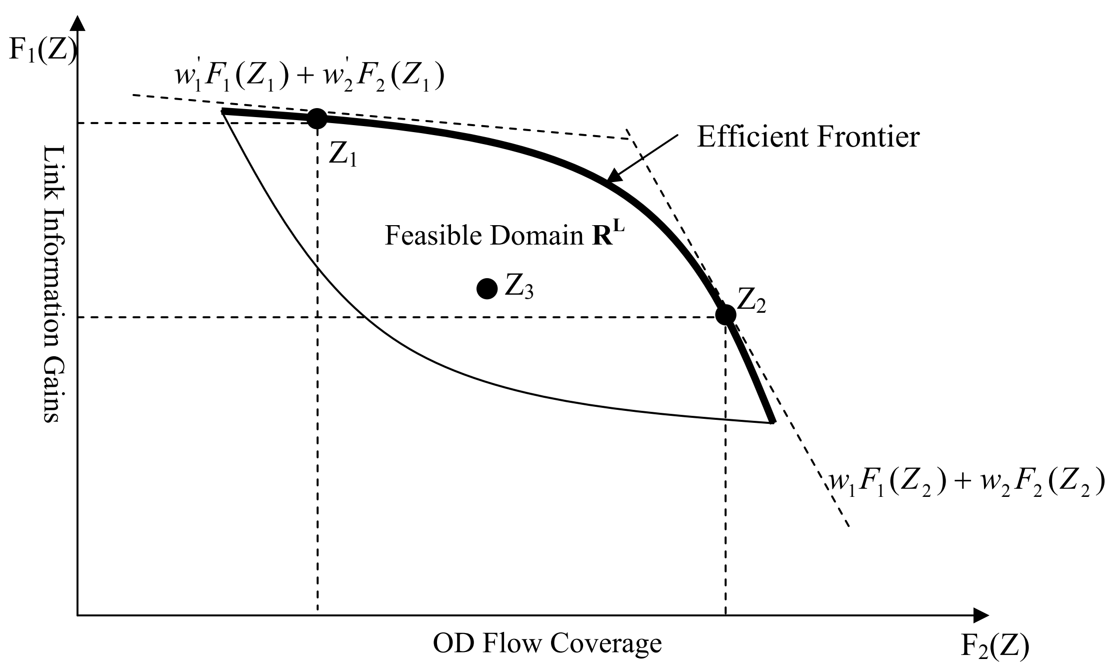


Figure 3-3 Graphic Definition of the Pareto Optimal

This theoretical discussion is intended primarily to frame the analysis conducted in the next section, and provide a conceptual framework for contemplating and understanding the sensor location problem of interest.

3.3 Model Formulation

This section presents methodological approaches in the context of dynamic traffic assignment to two variants, without and with budgetary constraint, of the sensor location problem. The first methodology is focused on solving the sensor location problem with an unlimited number of sensors (without budgetary constraint). The second methodology is focused on solving the sensor location problem with a given number of sensors (with budgetary constraint).

3.3.1 Notations and Problem Definition

Let $G=(V,A)$ represents a directed traffic network, with the set V of nodes and the set A of edges with the size $|A|=m$. Defines:

N : set of zones, consisting of n zones, size of set $|N|=n$

I : set of origin zones, consisting of n zones

J : set of destination zones, consisting of n zones

A : set of links, consisting of n_{LK} links, size of set $|A|=n_{LK}$

W : set of O-D pairs, size of set $|W|=n_{OD}$

L : set of links with measurements, size of set $|L|=n_m$

R_r^t : set of paths connected O-D pair r at departure time t

K : set of paths of the network, size of set $|K| = k$, $k = |R_1^1 \cup R_2^1 \cup \cdots \cup R_r^t \cup \cdots|$

a : subscript for link in network, $a \in A$

w : subscript for OD pair in network, $w \in W$

i : subscript for origin zone in network, $i \in I$

j : subscript for destination zone in network, $j \in J$

P_{w_1,w_2}^t : *a priori* demand covariance between OD pair w_1 and w_2 at time t,

$P_{w_1,w_2}^t \in R$

$\mathbf{C}$: vector of measurements (n_l x 1)

$\mathbf{H}$: mapping matrix (n_l x n_{OD}) mapping the demand flow to link counts

$\mathbf{D}$: demand vector, consisting of n_{OD} entries $d(i, j) \in \mathbf{D}$

$\hat{\mathbf{D}}(-)$: a *priori* estimated demand vector, consisting of n_{OD} entries $\hat{d}_{(i,j)}(-) \in \hat{\mathbf{D}}(-)$

$\hat{\mathbf{D}}(+)$: a *posteriori* estimated demand vector, $\hat{d}_{(i,j)}(+) \in \hat{\mathbf{D}}(+)$

$\widetilde{\mathbf{D}}(+)$: a *posteriori* estimated demand error matrix

$\widetilde{\mathbf{D}}(-)$: a *priori* estimated demand error matrix

$\mathbf{P}_{\hat{\mathbf{D}}}(-)$: a *priori* variance covariance matrix of the demand matrix

$\mathbf{P}_{\hat{\mathbf{D}}}(+)$: a *posteriori* variance covariance matrix of the demand matrix

d_w^τ : the ground truth O-D trips of O-D pair w at departure time τ

$\hat{d}_w^\tau(-)$: *a priori* estimated demand of O-D pair w at departure time τ

$\hat{d}_w^\tau(+)$: the *posteriori* estimated demand of O-D pair w at departure time τ

$Cov(i, j)(-)$: a *priori* variance covariance matrix of the demand matrix

$Cov(i, j)(+)$:a *posteriori* variance covariance matrix of the demand matrix

$k_{l,w}^{t}$: Kalman gain of link l from O-D pair w at time t

$h_{l,w}^{\tau,t}$: Assignment proportion of O-D pair w on link l departed at time τ at observation time interval t

T: set of all departure time intervals in the estimation period.

ε : vector of random noise quantities $\sim N(0, \mathbf{R})$ corrupting the measurements

3.3.2 Unlimited Network Sensors

Yang *et al*. (1998) formulated a binary integer program to determine the minimum number of sensor locations required to satisfy an OD covering rule for a road network with a given priori OD matrix and path selection.

$$Minimize \sum_{a \in A} z_a$$

$$subject\ to:$$

$$\sum_{a \in A} \delta_{aw} z_a \geq 1, \quad w \in W$$

$$z_a = 0,1,\ a \in A$$

Where $z_a = 1$ if a sensor is located on link *a* and zero otherwise; $\delta_{aw} = 1$ if some trips between O-D pair w, cross link $a \in A$ and zero otherwise. It can be shown that the resultant sensor location solution satisfies the OD covering rule and that selected links will be independent. A large network containing many OD zones and a significant number of links may be difficult to solve with this formulation. A heuristic used to solve the proposed formulation might only find a set of feasible or sub-optimal solutions instead of the optimal set. This is due to the trade-off between computation time and solution quality. In addition, Yang's model (1998) is based on static traffic assignment and considers an OD pair covered once a sensor is located on a single link of the paths between that particular OD pair. In reality, the path set between OD pairs evolves with time of the day. Thus, this OD covering model does not provide a valid result in that not every OD pair is assuredly covered at all times through the day.

To account for sensor location problems on large scale networks with time varying flows (e.g. determined using Dynamic Traffic Assignment (DTA) methodology), a method is proposed that considers time varying path-determinant. This model will result in a set of sensor locations on the links along the paths covering a subset of OD pairs, which experience OD demand flows in excess of a minimum number of trips, ζ^{τ}, where ζ^{τ} is a threshold termed as a " *degree*" to the relevant OD pairs at any time interval. Note that sensor location problem is mainly determined by the route choice and traffic assignment. Consequently, the following binary integer program formulation of the Deterministic Optimal Sensor Location Problem (DOSLP-1) is presented, subject to the coverage of the OD pairs with flow beyond a predefined "*relevant degree*" ζ^{τ}.

$$DOSLP-1 \qquad Minimize \sum_{a \in A} z_a^{\tau}$$

$$subject\ to:$$

$$\sum_{a \in A} \delta_{aw}^{\tau} z_a^{\tau} \geq 1 \quad w \in W,\ where\ d_w^{\tau} \geq \zeta^{\tau}, \tau \in T$$

$$z_a^{\tau} = 0, 1, a \in A$$

$$\delta_{aw}^{\tau} = assignment \ \lfloor d_w \rfloor \ from \ DTA, a \in A, \ w \in W, \tau \in T$$

Where $z_a^\tau = 1$ if a sensor is located on link a during departure time τ and zero otherwise. $\delta_{aw}^\tau = 1$ if some trips of OD pair w with departure time τ pass over link $a \in A$, and 0 otherwise. T is the planning horizon for sensor data collection.

3.3.3 Limited Network Sensors

Although a sensor network with full sensor coverage can infer all the O-D flows in a network, the mostly occurred situation to the transportation planners and decision makers is to deploy a given number of sensors in a large road network subject to the budgetary constraint. As aforementioned, this section examines the sensor location problem with a finite number of sensors being placed, which simultaneously consider maximizing the link information gains and O-D flow coverage under the assumption of recurrent traffic condition. Using the linear weighting method, the bi-objective optimization problem could be aggregated to a single objective optimization problem. The weights w_1, w_2 are determined by the decision maker's preference according to his/her experiences. The deterministic optimal sensor location problem (DOSLP) with limited sensor number is formulated as $DOSLP-2$.

$$DOSLP-2 \quad F(\mathbf{Z}) = Max\left\{ w_1 \sum_{t\in T}\sum_{l\in A}\sum_{w\in W}(k^t_{l,w} * z_l) + w_2 \sum_{\tau\le t\in T}\sum_{l\in A}\sum_{w\in W}\hat{d}^{\tau}_w(-) * I(h^{\tau,t}_{l,w}) * z_l \right\} \quad (3.27)$$

$s.t:$

$$w_1 + w_2 = 1 \quad (3.28)$$

$$\sum_{l\in A} z_l \le L \quad (3.29)$$

$$k^t_{l,w_1} = \frac{\sum_{\tau\le t}\sum_{w_2\in W} P^t_{w_1,w_2} * h^{\tau,t}_{l,w_2}}{\sum_{\tau\le t}\sum_{w_2\in W} h^{\tau,t}_{w_1,l} * P^t_{w_1,w_2} * h^{\tau,t}_{w_2,l} + r_l}, \qquad \forall w_1, w_2 \in W \quad (3.30)$$

$$I(h^{\tau,t}_{l,w}) = \begin{cases} 1, & if\ h^{\tau,t}_{l,w} > 0,\ \forall l, w, w\ departed\ at\ \tau\ has\ not\ been\ \text{cov}ered\ by\ t \\ 0, & Otherwise \end{cases} \quad (3.31)$$

$$h^{\tau,t}_{l,w} = F\left[l, \hat{d}^{\tau}_w(-), \forall w, l \in A, \tau \le t\right] \quad (3.32)$$

$$0 \le \tau \le t \le T \quad (3.33)$$

$$\hat{d}^{\tau}_w(+), \hat{d}^{\tau}_w(-) \ge 0, k^{\tau}_{l,w}, P^t_{w_1,w_2} \in R,\ w, w_1, w_2 \in W \quad (3.34)$$

$$z_l = 0,1,\ l \in A \quad (3.35)$$

The objective function (3.27) is composed of link information gains and O-D flow coverage. Constraint (3.28) shows that the summation of weights of all objectives must be 1. Constraint (3.29) indicates the total available sensors is L.

Constraint (3.30) is the information gain on link l brought by the measurement of O-D pair w_1 during observation interval t.

Constraint (3.31) is an indicator function that assures an O-D flow departed from its origin during time interval τ is counted only once in time t.

Constraint (3.32) is the assignment matrix coming from DYNASMART-P simulation result.

Constraint (3.33) specifies the simulation horizon. Constraint (3.34) is the non-negative constraint to the state variable.

Constraint (3.35) indicates a binary decision variable. If the sensor location is selected, the decision variable is 1; otherwise it is 0.

Since the two objectives are valued in different measurement scales, each objective of objective function (3.27) must be normalized before its weight is applied, as follows:

$$F(\mathbf{Z}) = Max\left\{ w_1 \sum_{t\in T}\sum_{l\in A}\sum_{w\in W}\left(\frac{k_{l,w}^{t}}{k_{\max}} * z_l\right) + w_2 \sum_{\tau\le t\in T}\sum_{l\in A}\sum_{w\in W}\frac{\hat{d}_w^{\tau}(-)}{\hat{d}_{\max}(-)} * I(h_{l,w}^{\tau,t}) * z_l \right\} \tag{3.36}$$

where $k_{\max} = Max\{k_{l,w}^{t}, \forall l, w, t\}$, is the maximal link information gain across the planning horizon. $\hat{d}_{\max}(-) = Max\{\hat{d}_w^{\tau}(-), \forall w, \tau\}$, is the maximal *a priori* O-D demand. In matrix form, Eq.3.36 reduces to

$$F(\mathbf{Z}) = Max\left\{ w_1 \sum_{t\in T}\left(\frac{\mathbf{K^t}}{k_{\max}} \cdot \mathbf{Z}\right) + w_2 \sum_{t\in T}\left(\frac{\mathbf{D^t}}{d_{\max}} \cdot \mathbf{I^t} \cdot \mathbf{Z}\right) \right\} \tag{3.37}$$

where $\mathbf{Z}$ is a $(n_A * 1)$ vector, $\mathbf{K^t}$ is a $(n_{OD} * n_A)$ link gain matrix of contributions by the sensors to the OD pairs during interval t, $\mathbf{D^t}$ is a $(n_{OD} * 1)^T$ vector during interval t, $\mathbf{I^t}$ is a $(n_{OD} * n_A)$ matrix during interval t.

An important issue about model DOSLP-2 is that the measurements did not play any role in the proposed model. This feature facilitates the evaluation of the selected locations especially to a large-scale traffic network.

3.3.4 Model Robustness

In order to assess the impact of different sensor location strategies in conjunction with the O-D demand estimator error reduction, the root mean squared error (RMSE) of the O-D demand will be calculated in order to check the quality of the estimated O-D matrix. The root mean squared error (RMSE) is simply the square root of the MSE.

Proposition: The proposed deterministic optimal sensor location model (DOSLP-2) always produces the minimal MSE across all other O-D estimators.

Proof: In statistics, the mean squared error (MSE) is defined as (Greene, 2000)

$$MSE(\theta \mid \hat{\theta}) = Var(\hat{\theta}) + E[(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T] \qquad (3.37)$$

As aforementioned, the GLS O-D demand estimator is unbiased; thus its MSE matrix is its covariance matrix. The MSE of the O-D estimator is

$$\mathbf{P}_{\hat{\mathbf{D}}}^{-1}(+) = \mathbf{P}_{\hat{\mathbf{D}}}^{-1}(-) + \mathbf{H}^{\mathbf{T}}\mathbf{R}^{-1}\mathbf{H}$$

Since $\mathbf{P}_{\hat{\mathbf{D}}}(-)$ is *a priori* variance covariance matrix of the demand matrix and the objective of the DOSLP-2 model is implicitly minimizing $P_{\hat{D}}(+)$, the MSE that based on the proposed models thus is the minimal statistics inference across all other estimators. This completes the proof □

3.4 Solution Procedure

The proposed models are computationally intensive. Model DOSLP-1 is a binary integer programming model, and the Branch-and-bound (BnB) can be used to solve this kind of problem. BnB is a problem solving strategy that is commonly used in solving computationally intensive integer programs. Due to its adaptability, BnB has been used in a variety of search algorithms, such as best-first search and depth-first search, as well as others.

Model DOSLP-2 is non-convex. Thus a global optimal solution is not guaranteed to exist. The solution procedure is formulated as a bi-level stochastic integer programming. The upper level is seeking the potential locations according to some selection rules, while at a lower level, the selected locations are evaluated using the results simulated by running user equilibrium from DYNASMART-P (Peeta & Mahmassani 1995).

3.4.1 Unlimited Network Sensors

Algorithm 3.1 illustrates the solution procedure for model DOSLP-1 based on Branch-and-Bound methodology.

Algorithm 3.1

Step 0: Run DYNASMART-P (Mahmassani et. al 2000) with *a priori* OD demand loaded to get $\delta_{aw}^{\tau}, a \in A,\ w \in W, \tau \in T, \tau = \tau_0\ , \zeta^{\tau} = \zeta_0^{\tau_0}$

Step 1: If $\tau < T$, filter out those OD pairs whose flow less than ζ^{τ}. Run Branch-and-Bound procedure to solve the binary integer model to obtain the solution set z^{τ} of DOSLP -1 during observation time interval τ. Otherwise if $\tau \geq T$, $Z = \bigcup_{\tau \in T} \{z_a^{\tau}\}$, Stop.

Step 2: Set $\tau = \tau + 1$, ζ^{τ} to satisfy the OD coverage percentage in time interval τ; go to step 1.

3.4.2 Limited Network Sensors

The major difficulty to solve DOSLP-2 is associated with the calculation of the Kalman gain matrix, because matrix inversion occurs at each time interval. The computational intensity is especially noticeable in a large-scale network. The sequential algorithm by Chui & Chen (1991) has been designed to avoid direct computation of the inversion of the matrix, $\mathbf{HP}_{\hat{\mathbf{D}}}\mathbf{(-)H}^{\mathbf{T}} + \mathbf{R}$ by assuming independence of the link measurement errors.

DOSLP problem is a $\mathcal{NP}$-hard problem. The likelihood optimization in (3.27) is quite formidable, and we are not aware of any method for computing the global maximum except by a brute force examination of each possible solution that select n_m links every time from the network $G(V, A)$, calculate the total link gains each time and then select the locations with the largest link gains. However the search space is the combination of n_m links from total n_{LK} links, namely $\binom{n_{LK}}{n_m} = \frac{n_{LK}!}{n_m!(n_{LK} - n_m)!}$ which results in a non-polynomial computational time. This explosion of the search space precludes the brute force approach in all but very small networks. It is imperative to develop an efficient and tractable solution procedure to find an optimal set of senor locations for large scale networks.

While determining the global optimal solution is prohibitive in most cases, a suboptimal algorithm based on bi-level programming technique is used in this study to solve the proposed sensor location problem. The proposed algorithm is a recursive selection process. In the upper level, a Greedy Randomized Adaptive Search Procedure (GRASP), as a combinatorial optimization algorithm, is developed to find feasible solutions through reducing the effective size of feasible solution space and exploring the space efficiently. In the lower level, the selected locations from the upper level are evaluated using the simulated results, e.g. assignment matrix, link information gains, etc. through running user equilibrium (UE) of DYNASMART-P. Details about user equilibrium (UE) and system optimization (SO) can be found in Peeta and Mahmassani (1995).

3.4.2.1 Hybrid Greedy Randomized Adaptive Search Procedure (HGRASP)

Greedy Randomized Adaptive Search Procedure is a multi-start or iterative sampling method (Lin & Kernighan, 1973, Feo & Resende 1995, Festa & Resende 2001, Pitsoulis and Resende 2001), with each GRASP iteration composed of two phases, a solution construction phase, where a randomized greedy solution is constructed, and a solution improvement (local search) phase, which starts at the constructed solution and applies iterative improvement until a locally optimal solution is found. The procedure of the HGRASP procedure for the proposed sensor location problem is as follows:

Algorithm 3.2

Step 0 (Initialization): Set $F^* = F(\mathbf{Z}^*) = -\infty$, where $\mathbf{Z}^*$ is the solution vector representing the best locations found so far.

Step 1 (Construction & Searching): Repeat if GRASP stopping criterion is not satisfied.

(a). Construct a greedy randomized solution $\mathbf{Z}$

(b). Local Search (Tabu Search): finding local optimal vector Z' in the neighborhood $N(\mathbf{Z})$

(c). Update Solution: if $F(\mathbf{Z}') > F^*$, $let\ F^* = F(\mathbf{Z}')\ and\ \mathbf{Z}^* = \mathbf{Z}'$, go to step 1

Step 2 (Best Solution Found): Return the best locations found $\mathbf{Z}^*$

In the construction phase, the candidate elements ranked with respect to a greedy function, which measures the benefit of choosing each element, are randomly selected one by one at each time. Pitsoulis and Resende (2001) summarized different random element selection methods to build a list of best candidates but not necessarily the top candidates in each HGRASP iteration. The list is called restricted candidate list (RCL). This selection technique enables the heuristic to diversify the exploration in the search space. This selection technique enables the heuristic to diversify the exploration in the search space. In this study, a randomly generated $\alpha \in UNIF[0,1]$ value coupled with an adaptive greedy function were used to build the RCL at each HGRASP iteration. Below is the procedure followed in the construction phase:

Construct a greedy randomized solution $\mathbf{Z}$

Step 0 (Initialization): $Set\ \mathbf{Z} = \{\ \}$

Step 1 (Construction): Repeat until the total elements in set $\mathbf{Z}$ equal to the number of sensors n_m

(a). $\hat{c}_{\max} = Max\left\{\hat{c}^l \mid \hat{c}^l = \sum_{\tau \le t} \sum_{w \in W} (h_{l,w}^{\tau,t} * \hat{d}_w^{\tau}(-)), l \in A\right\}$, where $\hat{c}_{\max}$ is the maximal link flow across the entire planning horizon T

(b). $RCL = \{l \in A \mid \hat{c}^l \ge \rho * \hat{c}_{\max}\}$, where $\rho \in [0,1]$ is a scalar.

(c). Pick l at random from RCL, while $l \notin \{L_k \mid L_k \in R_r^t(l_z), \forall l_z \in Z, t \le T\}$

(d). $Z = Z \cup \{l\}, A = A \setminus \{l\}$

Step 2: Return the solution set Z

$R_r^t(l_z)$ is a set of paths that traverse link l_z connecting O-D pair r during time interval t; L_k denotes the set of links comprising path k. Step 1(c) shows that the candidate link l cannot be on any path k that traversed the counting stations on those selected links in set Z. The inherent idea in step 1(c) is to select links with large information gains while keeping the rank of assignment matrix $\mathbf{H}$ full. By keeping the selected links independent, the procedure is trying to acquire more information. It should be noted that the measurements from those locations between which there are no intermediate intersections or entry/exit ramps are highly correlated with each other and will not contribute new traffic information. Step (1(c)) rules out the aforementioned possible sensor sites that may be located on the upstream or downstream points or do not have any entry or exit points between them.

Generally speaking, the solutions from the HGRASP construction phase are not usually locally optimal, thus a local search procedure needs to be employed to exploit the neighborhood $N(\mathbf{Z})$ of solution $\mathbf{Z}$ in each HGRASP iteration. Tabu Search, introduced by Glover (1987), is a metaheuristic method for intelligent problem solving (Glover and Laguna, 1993). The power and essential feature of Tabu search is the systematic adaptive use memory to record historical information for guiding the search process. The use of the short term memory strategy (Tabu list) helps to forbid (or tabu) the moves in pre-defined iterations (Tabu tenure) that might revisit recently visited solutions. The Tabu tenure helps to prevent cycling. A move applied to a solution will be a Tabu move if the Tabu conditions identified by the attributes (i.e. sensor location) are satisfied. However, with the aspiration level conditions, the Tabu status can be overruled if some Tabu solution has attractive results.

The following describes the local search procedure for the proposed sensor location problem. Recency-based Tabu memory functions were used to identify the starting and ending iterations of an attribute during the time that the attribute is Tabu-active. A dynamic neighborhood structure was employed in this study.

Local Search: finding local optimal vector $\mathbf{Z}'$ *in the neighborhood* $N(\mathbf{Z})$

Step 0 (Initialization): *Set* $k = 0$, empty the tabu list

Step 1: Repeat until the stopping criterion is satisfied

(a) (*Drop Move*). Randomly choose a location $x \in Z$

(b) (*Add Move*). Set $k = k + 1$, $N(x,k)$ is the path set of the neighborhood of x at step k, where $N(x,k) = \{l \mid l \in R^t_{(i,j)}(x),\ 0 \le t \le T\}$.

A logit formulation is used to determine the selection probability, which let all of the links likely be selected while those links with larger flows have higher likelihood to be selected.

$$P_l = \frac{e^{\alpha \cdot \hat{c}^l}}{\sum_{i \in N(x,k)} e^{\alpha \cdot \hat{c}^i}} \tag{3.38}$$

where P_l is the probability for choosing link l

$\hat{c}^l = \sum_{\tau \le t} \sum_{w \in W} (h_{l,w}^{\tau,t} * \hat{d}_w^{\tau}(-)), l \in N(x,k)$ is the summation of simulated link flows on link l during planning horizon T

α is a scaling parameter

Scanning the Tabu list, if the selected link l is not on the list or if the selected link l is on the list, but aspiration criteria is met, put this link l at the bottom of the list. Otherwise, ignore this link and choose another link l', Set $Z' = (Z/\{x\}) \cup \{l\}$

(c) (*Update*). If $F(\mathbf{Z}') > F(\mathbf{Z})$, *Set* $\mathbf{Z} = \mathbf{Z}'$, $F(\mathbf{Z}) = F(\mathbf{Z}')$, update the Tabu list and aspiration conditions.

Step 2: Return the local optimal solution set $\mathbf{Z}$

The proposed HGRASP-DTA heuristic starts from a set of initially independent locations, and iteratively explores the neighborhoods of current solution till the stopping criteria satisfied. Thus the locations of the final result could be either independent or dependent that depends on decision makers' preference to reducing demand uncertainty or increasing O-D flow coverage. The HGRASP-DTA flow chart of the proposed process is shown in Figure 3-4.

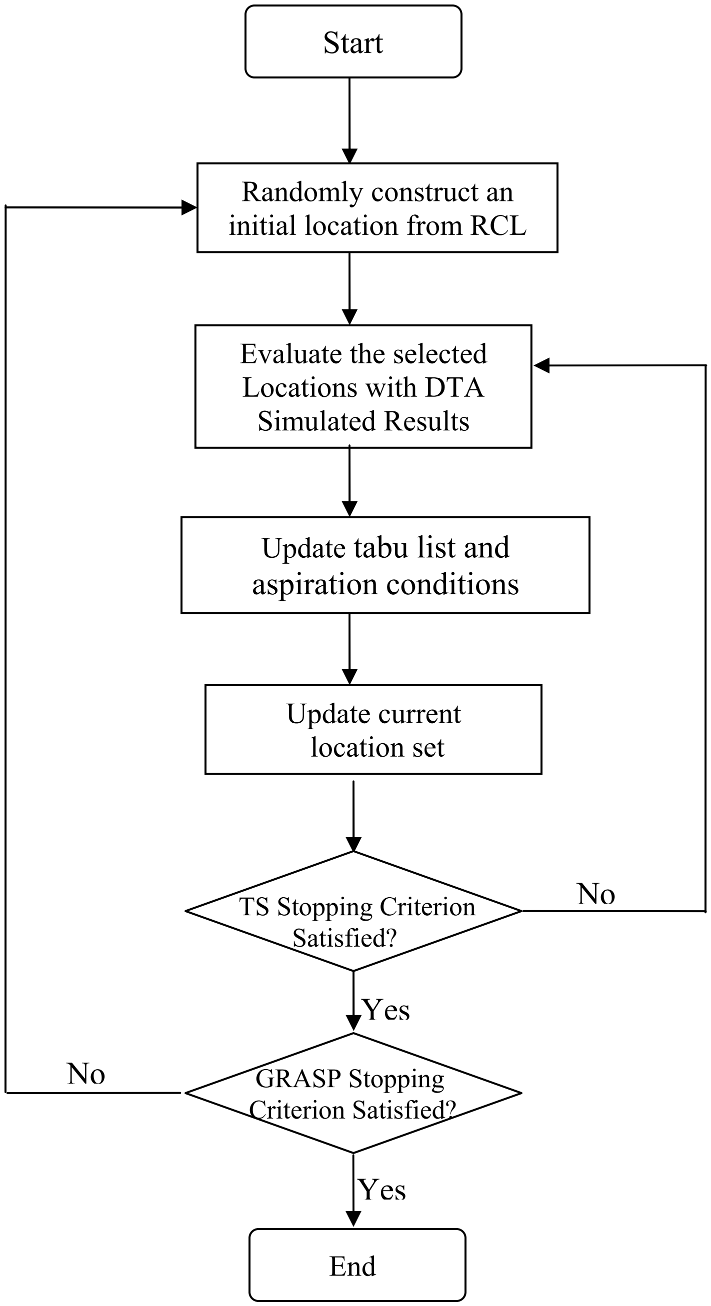
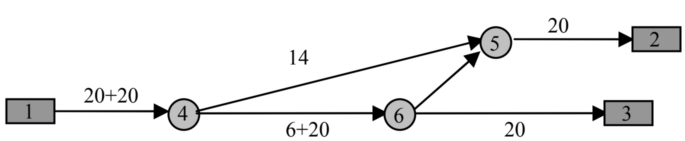


Figure 3-4 Hybrid GRASP-DTA bi-level solution procedure

3.5 Numerical Illustration

A series of examples based on a small 6-node network is used to demonstrate the proposed methodology. In order to facilitate ability to compare the results of this research to the recent results of Zhou and List (2006), the same example network was used.

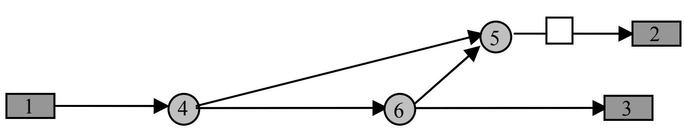
The first example is a single point sensor location, according to the set-up in figure 3-5. OD pair 1 is from node 1 to node 2 and OD pair 2 is from node 1 to node 3; OD pair 1 has two routes; and 70% of the flow travels along path $\{1 \quad 4 \quad 5 \quad 2\}$ while the remaining 30% of the flow travels along path $\{1 \quad 4 \quad 6 \quad 5 \quad 2\}$. Both OD pairs have a flow volume of 20 units. Assume $P_{\hat{D}}(-) = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$, meaning that OD pair 1 has a larger *a priori* variance than OD pair 2. The standard deviation of the measurement error for a sensor is assumed to be 5% of the corresponding true flow volume.



$$P_D = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$$

$$K = 0$$

Base Case

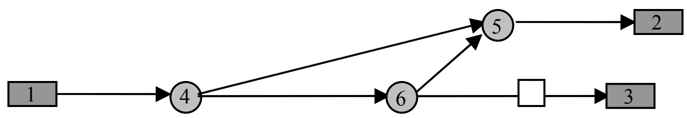


(a) One sensor for OD pair (1->2)

$$R = 1 \quad H = [1 \quad 0]$$

$$K = \begin{bmatrix} 0.8 & 0 \end{bmatrix}^T$$

$$Gain = \sum_{w=1}^{2} K_a = 0.8$$

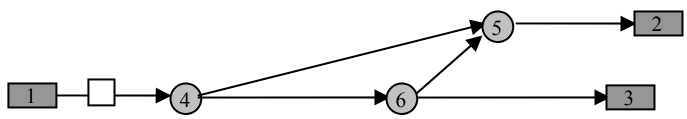


$R = 1 \quad H = [0 \quad 1]$

$K = \left[0 \quad 0.5\right]^T$

$Gain = \sum_{w=1}^{2} K_a = 0.5$

(b) One sensor for OD pair (1->3)



$R = 2^2 \quad H = \left[1 \quad 1\right]$

$K = \left[0.6667 \quad 0.1667\right]^T$

$Gain = \sum_{w=1}^{2} K_a = 0.8334$

(c) One sensor for both OD pair

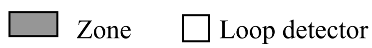


Figure 3-5 Examples of Single Point Sensor Locations

Figure 3-5 illustrates single point sensor locations in the network. Sensor in (a) covers O-D pair 1 with larger variance and produces larger gain than that in (b). Since the sensor in (c) covers both O-D pairs and intercepts more OD flows in these three scenarios, it gets the largest gain through the observation counts even though it has larger measurement error than that in (a) and (b). If the error in (c) is reduced to 1, as in (a) and (b), it has R=1, $K = \left[0.6667 \quad 0.1667\right]^T$, and $Gain = 0.8337$, producing larger information gain.

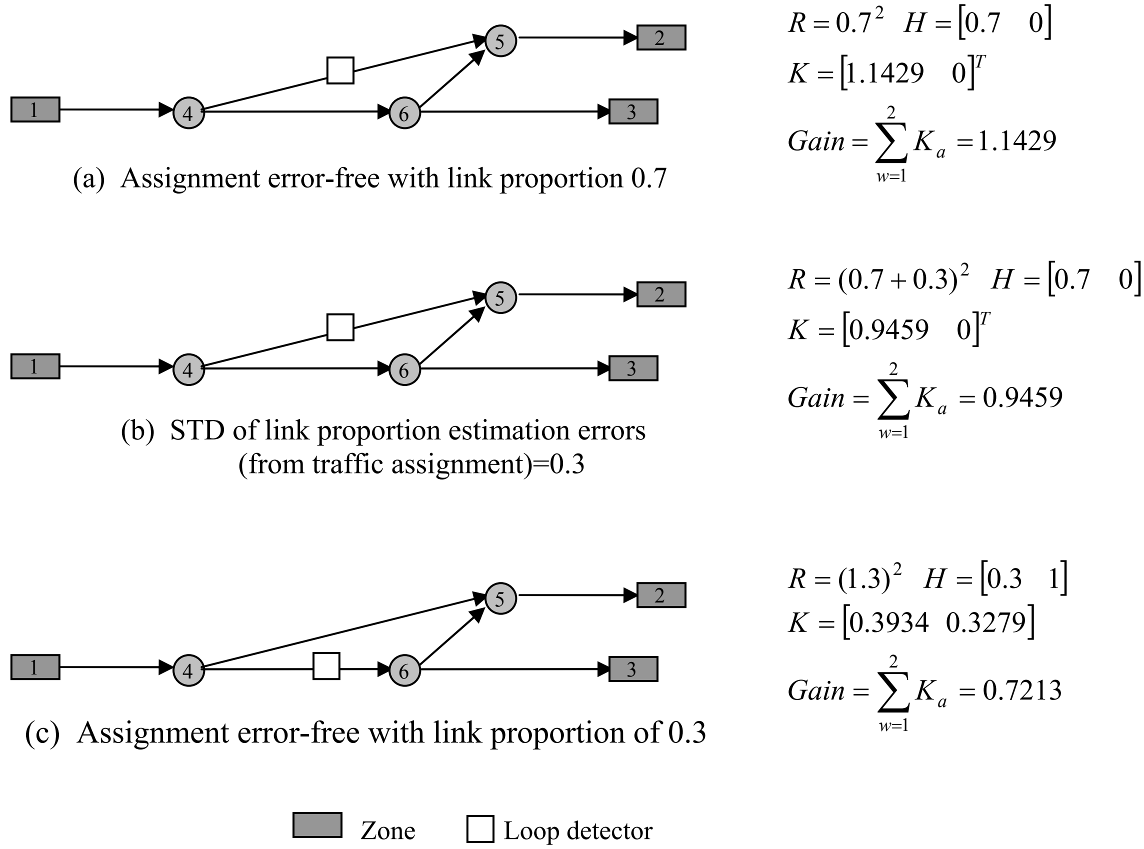
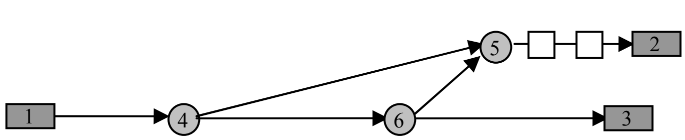


Figure 3-6 Examples of Single Point Sensor Locations with Route Choice

Figure 3-6 shows the examples of single sensor locations with route choice. Scenario (a) shows an error free link proportion estimate and the measurement error proportional to the link flow scenario. The gain in scenario (a) is 1.1429, which is greater than all the scenarios in Figure 3-5. This indicates that the measurement error can reduce the link information gain. Scenario (b) shows that the link proportion estimation error could also reduce the information gains. Although the sensor in Scenario (c) covers both OD pairs, it still cannot produce the largest information gain because of the largest measurement error in the three scenarios. Even when the measurement error is reduced to 1, the gain matrix is $K = [0.5085 \quad 0.4237]^T$ and the gain is 0.9322.

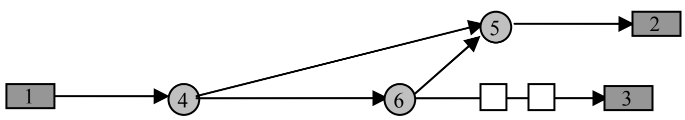


(a) Two uncorrelated sensors for OD pair (1->2)

$$R = I \quad H = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

$$K = \begin{bmatrix} 0.4444 & 0.4444 \\ 0 & 0 \end{bmatrix}$$

$$Gain = \sum_{a=1}^{2} K_a = 0.8888$$

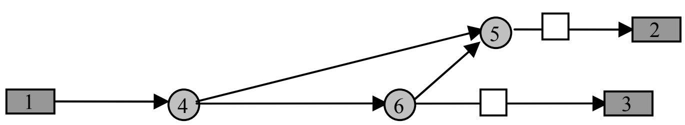


(b) Two uncorrelated sensors for OD pair (1->3)

$$R = I \quad H = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$$

$$K = \begin{bmatrix} 0 & 0 \\ 0.3333 & 0.3333 \end{bmatrix}$$

$$Gain = \sum_{a=1}^{2} K_a = 0.6666$$

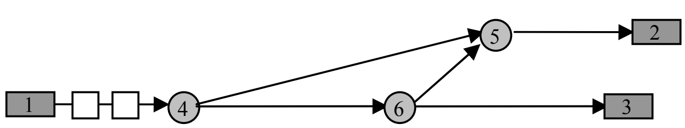


(c) Two uncorrelated sensors for both OD pairs

$$R = I \quad H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$K = \begin{bmatrix} 0.8 & 0 \\ 0 & 0.5 \end{bmatrix}$$

$$Gain = \sum_{a=1}^{2} K_a = 1.3$$

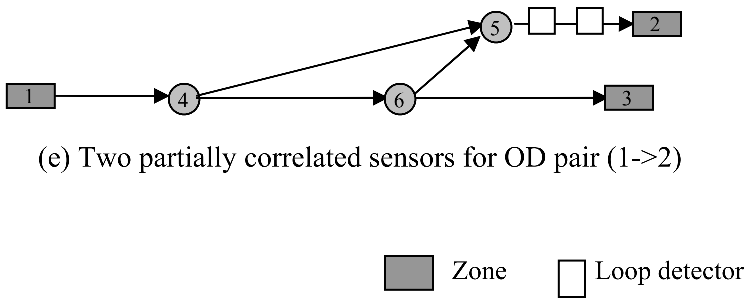


(d) Two uncorrelated sensors for both OD pairs

$$R = I \quad H = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$K = \begin{bmatrix} 0.3636 & 0.3636 \\ 0.0909 & 0.0909 \end{bmatrix}$$

$$Gain = \sum_{a=1}^{2} K_a = 0.909$$



(e) Two partially correlated sensors for OD pair (1->2)

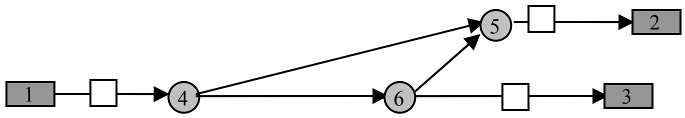
$$R = \begin{bmatrix} 1 & 0.25 \\ 0.25 & 1 \end{bmatrix} \quad H = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

$$K = \begin{bmatrix} 0.4324 & 0.4324 \\ 0 & 0 \end{bmatrix}$$

$$Gain = \sum_{a=1}^{2} K_a = 0.8648$$

Figure 3-7 Examples of Two Point Sensor Locations

Figure 3-7 shows examples of two sensor locations. Scenario (a) covers O-D pair 1, Scenario (b) covers O-D pair 2, Scenario (c) and Scenario (d) covers both O-D pairs, Scenario (e) covers O-D pair 1 but the two sensors have measurement error correlation between them. As expected, scenario (c) collected the larger gains than other scenarios since it covers both OD pairs and the two sensors are independent with each other. Although scenario (d) covers both OD pairs as well, the information gain is smaller than scenario (c) due to the linear dependence of the two observations. Comparing (a) and (e), the correlation of measurement errors made some reduction of the information gain.

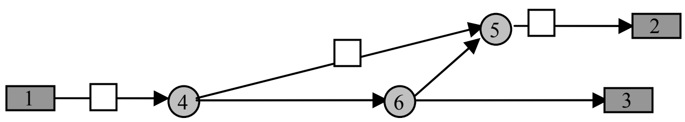


(a) Three uncorrelated sensors for both OD pairs

$$R = 1.5I \qquad H = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$K = \begin{bmatrix} 0.3419 & 0.4786 & -0.1368 \\ 0.1880 & -0.1368 & 0.3248 \end{bmatrix}$$

$$Gain = \sum_{a=1}^{3} K_a = 1.3333$$

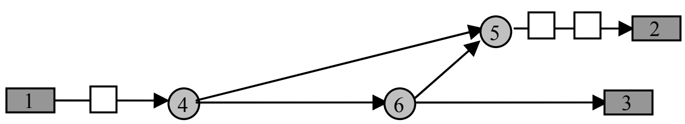


(b) Three uncorrelated sensors for both OD pairs

$$R = 1.5I \qquad H = \begin{bmatrix} 1 & 1 \\ 0.7 & 0 \\ 1 & 0 \end{bmatrix}$$

$$K = \begin{bmatrix} 0.2434 & 0.2840 & 0.4057 \\ 0.3026 & -0.1136 & -0.1623 \end{bmatrix}$$

$$Gain = \sum_{a=1}^{3} K_a = 1.2357$$

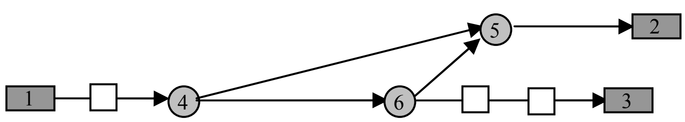


(c) Three uncorrelated sensors for both OD pairs

$$R = 1.5I \qquad H = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}$$

$$K = \begin{bmatrix} 0.2017 & 0.3361 & 0.3361 \\ 0.3193 & -0.1345 & -0.1345 \end{bmatrix}$$

$$Gain = \sum_{a=1}^{3} K_a = 1.1932$$

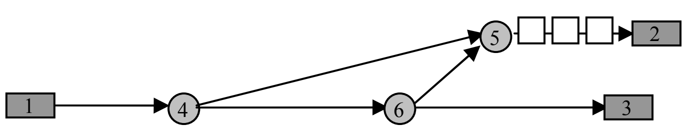


(d) Three uncorrelated sensors for both OD pairs

$$R = 1.5I \qquad H = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

$$K = \begin{bmatrix} 0.6747 & -0.1928 & -0.1928 \\ 0.0723 & 0.2651 & 0.2651 \end{bmatrix}$$

$$Gain = \sum_{a=1}^{3} K_a = 1.2772$$



(e) Three uncorrelated sensors for OD pair (1->2)

$$R = 1.5I \qquad H = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}$$

$$K = \begin{bmatrix} 0.2963 & 0.2963 & 0.2963 \\ 0 & 0 & 0 \end{bmatrix}$$

$$Gain = \sum_{a=1}^{3} K_a = 0.8889$$

Figure 3-8 Examples of Three Point Sensor Locations

Figure 3-8 shows the examples of three sensor locations. The scenario (e) collected the least information gain since the three sensors covered only one OD pair while other scenarios cover both OD pairs. Scenario (a) produces the best gain because of the link independence of the sensor data.

An interesting finding from above examples is that more sensors do not always result in more information gain. Scenario (c) with 2 sensors in figure 3-6 ($Gain = 1.3$) has larger gains than most scenarios in figure 3-7. Even if the two cases have the same measurement errors, the scenario (e) in figure 3-8 covering 1 OD pair, has $K = \begin{bmatrix} 0.3077 & 0.3077 & 0.3077 \\ 0 & 0 & 0 \end{bmatrix}$, $Gain = 0.9231$, which is less than that in scenario (c) in figure 3-6.

Under the assumption that the simulated assignment matrix reflects the actual route choice in the proposed sensor location problem, it can be proved by the general linear regression that only if the assignment matrix $\mathbf{H}$ has full rank, the OD demand estimator $\hat{\mathbf{D}}(-)$ is the best linear unbiased estimator (BLUE). The gain matrix was derived based on the BLUE assumption which explained the reason why the independent sensor data always produced the largest gains. The following observations are made from the aforementioned example results. In order to maximize the information gains, (1) the sensors need to be located on the links that can intercept the most OD flows; (2) the sensor observation data should be linearly independent; (3) more sensors do not necessarily mean larger information gains; and (4) the lower the measurement error, the more gains the system could attain.

3.6 Summary

This chapter presents the sensor location problem in two different scenarios, without and with budgetary constraints. In the first scenario, the senor location problem is viewed as an O-D covering problem under dynamic traffic assignment. In the second scenario, a Kalman filtering based model is presented to explore time-dependent maximal information gains and O-D demand coverage across all the links in the network. The solution procedure is formulated as a bi-level stochastic integer programming. The upper level is seeking the potential locations according to some selection rules, while at the lower level, the selected locations are evaluated using the simulated results by running user equilibrium of DYNASMART - P. A hybrid greedy randomized adaptive search heuristics is developed for finding the near optimal sensor locations to circumvent the computational complexity of the proposed problem.

Recognizing the importance of sensor location and its relationship to the quality of OD demand estimation, this chapter built a connection between these two critical issues and considered demand estimation error based on Kalman filtering algorithm in the sensor location model formulation.

Chapter 4 A Two-Stage Stochastic Model for the Sensor Location Problem in a Large-Scale Network

4.1 Introduction

Uncertainty is one of the major factors that transportation system analysts and planners have to deal with in making transportation planning decisions. As part of network operational planning, transportation agencies may be in position to deploy a limited number of sensors in the network before any unpredictable events (e.g. incidents, weather, special events, etc). However, due to unavoidable day-to-day traffic demand evolutionary uncertainties and randomly occurring uncertain events which affect the traffic pattern in the network, there is a great need to develop a methodology to identify a valid sensor location strategy, which performs more robustly with regard to extreme cases. Network uncertainties, such as location, duration, and severity associated with most disasters limit the applicability of the deterministic model proposed in the last chapter under these situations. The nature of this design problem under uncertainty presents itself as a two-stage sequence of decisions. The first stage decision produces a strategic sensor location plan before observations of any random events, while the recourse function associated with the second stage denotes the expected cost of taking corrective actions to the first stage solution after the occurrence of the random events. Thus, the dynamic sensor location problem is formulated as a two-stage stochastic model with recourse in this chapter.

The proposed stochastic optimal sensor location model in this chapter is extended from the deterministic model presented in the previous chapter by accounting for network uncertainty in a mathematical program. The aim of the model is to determine valid sensor locations that may not be optimal for every possible realization of the un-anticipated cases, but perform more robustly with regard to extreme cases and thus is hedged against various network random occurrences.

The objective of this chapter is to provide a model that can be used to gain insight into the sensor location problem when both traffic dynamics and network uncertainty are accounted for in the model formulation.

The rest of this chapter is comprised of four sections. Section 4.2 introduces the potential problems for approaching the stochastic optimal sensor location problem. Section 4.3 proposes a model formulation for the stochastic sensor location problem and discusses an incident generation model under Poisson probability distribution assumption. Section 4.4 presents a model solution procedure. Section 4.5 summarizes the entire chapter.

4.2 Problem Statement

Sensor locations play a critical role in reducing the uncertainty of the estimated OD demand and consequently improve the quality of the predicted network OD demand as well as the system performance. It is generally recognized that incidents could lead to rapid deterioration of network performance. The stochastic model presented for the sensor location problem is used to evaluate the locations selected a priori, before incidents occurred, under different incident scenarios defined in terms of location, severity, and duration of the incident(s). By incorporating the impacts of randomly occurring incidents on the traffic pattern into the model formulation, this chapter extends the deterministic model to a stochastic model. It seeks to maximize the long-run average OD coverage and minimize the long-run average demand uncertainty in response to different incident realizations subject to a budget constraint.

One challenge inherent in the sensor location problem is the randomness of the events (i.e. incident location, duration, severity, etc) as well as the subsequent impact on the associated traffic pattern and traveler behavior dynamics. A stochastic programming framework is built to incorporate the uncertainty involved in this problem into the model formulation. Note that although the proposed stochastic model is general with regard to various types of uncertain events, this research is only focused on the impacts due to network traffic incidents. Without considering specific incidents, a set of sensor locations is identified in the first stage subject to budgetary constraints; a recourse decision is then made in the second stage based on the specific incident realizations in the network, which are consequently defined as random variables. Note that the location plans from the deterministic model can be used as the initial candidate locations in the stochastic model.

Another challenge in this sensor location problem is how to model the occurrence of highly uncertain incident events in the network. Chiu *et al.* (2001) assumed the occurrence of incidents on link a_n follows a Poisson process, and calculated the likelihood of n incidents occurring on link a_n using Bayesian statistical method. A method based on their incident generation model is adopted in this study to generate random incident realizations. The system uncertainties can conceptually be modeled by a scenario tree which describes system uncertainty evolution across all stages. The scenario tree with all possible incident scenarios (low severity s_L, medium severity s_M, high severity s_H on link L_i) on its leaves is used to produce robust sensor location strategies (Figure 4-1).

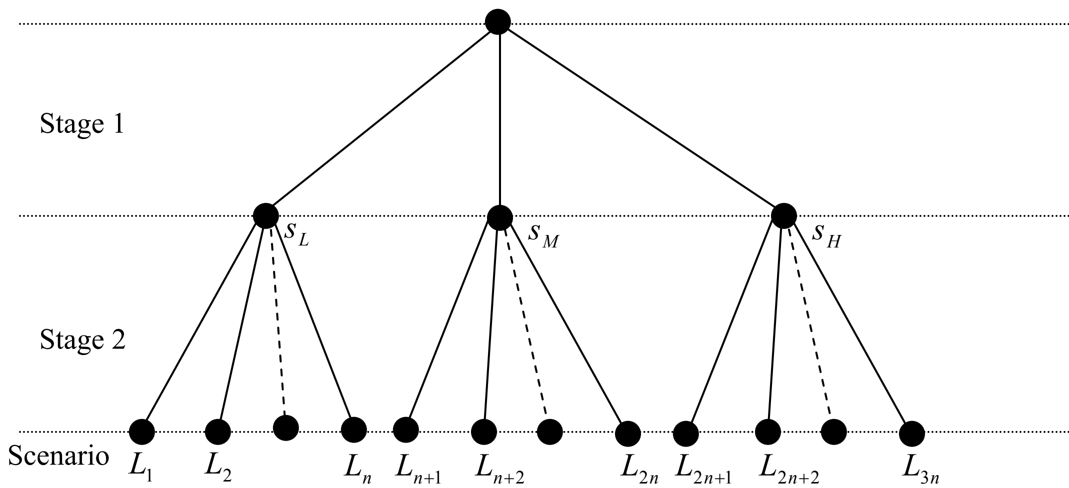


Figure 4-1 Scenario Tree for SOSLP with Scenarios on Leaves

4.3 Problem Formulation

Due to the intrinsic characteristics of the proposed Stochastic Optimal Sensor Location Problem (SOSLP), a bi-level stochastic mixed integer model framework is presented in this section. In the upper level, the traffic planner makes decisions on sensor placement in the network, to maximize the long run average OD flow coverage and minimize the expected uncertainty of the estimated OD demand subject to the budget limitation. In the lower level, the network users are assigned to the time-dependent user equilibrium or system optimization routes given the sensor locations determined from the upper level and are subject to the incident realizations. In this study, the network users are presumed to have full knowledge of the travel times over all the routes of interest. The traffic flow pattern is consequently assumed to be a user equilibrium (UE) which was initially introduced by Wardrop (1952), namely, that for each OD pair, at UE, the travel time on all used paths, no matter which combination of travel routes and departure times the traveler choose, are equal and less than or equal to the travel time that would be experienced by a single vehicle on any unused path. The UE constraints in the lower level of the proposed stochastic optimal sensor location problem (SOSLP) result in a mathematical program with equilibrium constraints (SMPEC) (Patriksson and Wynter, 1999).

The decision variables of the upper level are integer binary variables, which denote the sensor locations. The decision variables in the lower level are the assignment matrices induced by the time-dependent user equilibrium paths. The lower level equilibrium problem is based on the work of Peeta and Mahmassani (1995) and Chiu, Huynh and Mahmassani (2001), under the assumption that detours followed by impacted vehicles before reaching the incident scene would have only negligeable effect on the network performance. Given the small portion of the impacted vehicles to the total number of vehicles in a large-scale congested network, this assumption is reasonable. Due to the computational intensity and complexity of the assignment matrix in a large scale network, the time-dependent assignment matrix in this study is obtained from a simulation-based dynamic traffic assignment model software DYNASMART-P (Mahmassani *et al.* 2000), described in Chapter 2. The notation and problem definition are first introduced below before the model formulation.

4.3.1 Notation and Problem Definition

Let $G = (V, A)$ represents a directed traffic network, with the set V of nodes and the set A of edges with the size $|A| = m$. Defines:

N	Set of zones, consisting of n zones, size of set $\|N\| = n$
I	Set of origin zones, consisting of n zones
I_v^ω	Set of nodes where impacted vehicles receive reassignment under scenario ω
J	Set of destination zones, consisting of n zones
A	Set of links, consisting of n_{LK} links, size of set $\|A\| = n_{LK}$
W	Set of O-D pairs, size of set $\|W\| = n_{OD}$
L	Set of links with measurements, size of set $\|L\| = n_m$
P_ω	Probability of a random event ω (e.g. $P_\omega = P(\xi = \xi_\omega)$)
Ω	Set of all random events
ω	Random event ($\omega \in \Omega$) with respect to the probability space (Ω, P)
a	Subscript for link in network, $a \in A$
w	Subscript for OD pair in network, $w \in W$
i	Subscript for origin zone in network, $i \in I$
i_v^ω	Subscript for the node where impacted vehicles receive reassignment under scenario ω
j	Subscript for destination zone in network, $j \in J$
$OB(n)$	Set of outbound links from node n

$IB(n)$ Set of inbound links to node n

$C(n)$ Set of links terminating at node n

$O_n^{t,\omega}$ Number of the out of network vehicles from node n during time t under scenario ω

$I_n^{t,\omega}$ Number of the vehicles entering network from node n during time t under scenario ω

$E_n^{t,\omega}$ Number of vehicles generated at node n during time t under scenario ω

$m_a^{t,\omega}$ Total number of vehicles that enter link a during time t under scenario ω

$d_a^{t,\omega}$ Total number of vehicles that exit link a during time t under scenario ω

$x_a^{t,\omega}$ Number of vehicles on link a during time t under scenario ω

T Planning horizon

λ Objective function weight, $\lambda = \lambda_1, \lambda_2$, $0 \le \lambda \le 1$

τ Superscript denoting departure time interval, $0 \le \tau \le T$

V^{ω} Set of vehicles that are impacted by scenario ω

O^{ω} Set of vehicles that are not impacted by scenario ω

U Set of all vehicles. $U = V^{\omega} + O^{\omega}$

u Superscript for impacted/non-impacted vehicles, $u = v^{\omega}, o^{\omega}$

$\left[T_s^{\omega}, T_e^{\omega}\right]$ Incident ω duration

$R_{i,j}^{\tau,u,\omega}$ Set of paths connected origin i and destination j during departure time τ under scenario ω for impacted/non-impacted vehicles

K Set of paths of the network, size of set $|K| = k$, $k = \left|R_{1,2}^{1} \cup R_{1,3}^{1} \cup \cdots \cup R_{i,j}^{t} \cup \cdots\right|$

$k(u)$ Subscript for the paths of impacted/non-impacted vehicles in the network under scenario ω, $k(u) \in R_{i,j}^{\tau,u,\omega}$

$T_{i,j,k(u)}^{\tau,u,\omega}$ Experienced travel time of the network vehicles (impacted or non-impacted) leaving from i to destination j along path $k(u)$ at departure time τ under scenario ω

$\pi_{i_v^\omega,j,k(v)}^{\tau,v,\omega}$ Minimal travel time for the impacted vehicles rerouting from node i_v^ω to destination j along the path $k(v)$ under scenario ω

$\hat{d}_{i,j}^{\tau}(-)$ *a-priori* estimated OD demand from origin i to destination j at departure time τ

$\hat{r}_{i,j,k(v)}^{\tau,v,\omega}$ Number of impacted vehicles v^ω leaving from i to destination j along path $k(v)$ under scenario ω

C_a Link a capacity

$\widetilde{C}_a^{t,\omega}$ Reduced capacity of link a when incident occurred on it during time t under scenario ω

z_l Decision variable of the upper level problem

$$z_l = \begin{cases} 1, \textit{if the sensor located on link } l, & \forall l \in A \\ 0, \textit{Otherwise} \end{cases}$$

P_{w_1,w_2}^t a *priori* variance covariance of the between OD pair w_1 and w_2 at time t, $P_{w_1,w_2}^t \in R$

$\mathbf{P}_{\hat{\mathbf{D}}}(-)$: a *priori* variance covariance matrix of the demand matrix

$\mathbf{P}_{\hat{\mathbf{D}}}(+)$: a *posteriori* variance covariance matrix of the demand matrix

r_a : Standard deviation of the measurement error corrupting the measurements

$k_{a,w}^{t,\omega}$ Kalman gain of link a brought by O-D pair w at time t under scenario ω

$h_{a,w}^{\tau,t,\omega}$ Assignment proportion of O-D pair w on link a departed at time τ during observation time interval t under scenario ω

$h_{i,i_v^{\omega},j}^{\tau,t,\omega}$ Time-dependent node-path incidence indicator

$$h_{i,i_v^{\omega},j}^{\tau,t,\omega} = \begin{cases} 1, \textit{if vehicles generated from i to j at departure time } \tau \textit{ are leaving} \\ \quad \textit{from } i_v^{\omega} \textit{ at time t} \\ 0, \textit{otherwise} \end{cases}$$

H Mapping matrix (n_l x n_{OD}) mapping the demand flow to link counts

S Number of random events (incidents) in the network

F First stage objective function (e.g. $\min F = c^T x + \cdots$)

Q Second stage value function with random argument

Q Expected second stage recourse function

E Mathematical expectation operator

ξ Random vector (ξ^t , if indexed by time) with realizations as ξ (without boldface)

ξ_a^t Binary random variable

$$\xi_a^t = \begin{cases} 1, \textit{if incident occurs on link a during time t} \\ 0, \textit{otherwise} \end{cases}$$

σ^{ω} Incident severity described as percentage of link capacity reduction under scenario ω

Δ Length of a time interval

$I(h_{a,w}^{\tau,t,\omega})$ A binary indicator variable

$$I(h_{a,w}^{\tau,t,\omega}) = \begin{cases} 1, & if\ h_{a,w}^{\tau,t,\omega} > 0\ and\ w\ has\ not\ been\ \mathrm{cov}ered\ by\ time\ t,\ \ \forall a, w, \omega \\ 0, & Otherwise \end{cases}$$

4.3.2 Model Formulation

Akin to the deterministic model, the problem objectives are to maximize the expected OD coverage and minimize variation of the estimated OD matrix under different scenario $s \in S$. Eq (4-0) shows the relation between the demand a posteriori variance and the link information gains.

$$\mathbf{P}_{\hat{\mathbf{D}}}(+) = (\mathbf{I} - \mathbf{KH})\mathbf{P}_{\hat{\mathbf{D}}}(-) \qquad (4\text{-}0)$$

Apparently, the maximization of the link information gains $\mathbf{K}$ and the minimization of the uncertainty of the estimated OD demand $\mathbf{P}_{\hat{\mathbf{D}}}(+)$ are mathematically equivalent. The problem hence can be formulated as follows,

$$J = Max\{\Pr[F(\mathbf{Z}) + Q(\mathbf{H}(\omega, \mathbf{Z}))]\}$$

$$= Max\{F(\mathbf{Z}) + E_{\xi}(Q(\mathbf{H}(\omega, \mathbf{Z})))\} \qquad (4\text{-}0\text{-}1)$$

where $E_{\xi}(Q(\mathbf{H}(\omega, \mathbf{Z})))$ is referred to as a *recourse function*. In this study, there are no first-stage costs ($F(z) = 0$) in the objective function since the first-stage variable z_a is reflected by $\sum_{a \in A} z_a \le L$. The second stage value function can be formulated as follows, with the random arguments:

$$Q(\mathbf{H}(\omega, z)) = \lambda_1 \sum_{t \in T} \sum_{a \in A} \sum_{w \in W} (k_{w,a}^{t,\omega} * z_a) + \lambda_2 \sum_{t \in T} \sum_{a \in A} \sum_{w \in W} (\hat{d}_{i,j}^{\,t}(-) * I(h_{w,a}^{\tau,t,\omega}) * z_a) \qquad (4\text{-}0\text{-}2)$$

The weights λ_1, λ_2 reflect the decision maker's relative preference for OD coverage or link gains.

4.3.2.1 Stochastic Optimal Sensor Location Problem (SOSLP)

$$J = Max\left\{E_{\xi}(Q(\mathbf{H}(\omega,\mathbf{Z})))\right\} \qquad (4\text{-}1)$$

$$s.t:$$

$$\lambda_1 + \lambda_2 = 1 \qquad (4\text{-}2)$$

$$\sum_{a \in A} z_a \le L \qquad (4\text{-}3)$$

$$k_{w_1,a}^{t,\omega} = \frac{\sum_{\tau \le t}\sum_{w_2 \in W} P_{w_1,w_2}^{t} * h_{w_2,a}^{\tau,t,\omega}}{\sum_{\tau \le t}\sum_{w_2 \in W} h_{w_1,a}^{\tau,t,\omega} * P_{w_1,w_2}^{t} * h_{w_2,a}^{\tau,t,\omega} + r_a}, \qquad \forall w_1 \in W, \omega \qquad (4\text{-}4)$$

$$h_{w,a}^{\tau,t,\omega} = h_{i,j,a}^{\tau,t,\omega} = \psi_1(f_{\omega}(\hat{d}_{i,j}^{\tau}(-)), \qquad \forall w \in W, a, \tau t, i, j, \omega \qquad (4\text{-}5)$$

$$T_{i_v^{\omega},j,k(v)}^{\tau,v,\omega} \ge \pi_{i_v^{\omega},j,k(v)}^{\tau,v,\omega} \qquad \forall i_v^{\omega}, j, v, k(v), \tau, \omega \qquad (4\text{-}6)$$

$$\hat{r}_{i_v^{\omega},j,k(v)}^{\tau,v,\omega} * (T_{i_v^{\omega},j,k(v)}^{\tau,v,\omega} - \pi_{i_v^{\omega},j,k(v)}^{\tau,v,\omega}) = 0 \qquad \forall i_v^{\omega}, j, v, k(v), \tau, \omega \qquad (4\text{-}7)$$

$$\widetilde{C}_a^{t,\omega} = C_a * (1 - \sigma^{\omega} * \xi_a^{t}) \qquad \forall a, \omega, t \in \left[T_s^{\omega}, T_e^{\omega}\right] \qquad (4\text{-}8)$$

$$\sum_{OB(n)} d_a^{t,\omega} = \sum_{IB(n)} m_a^{t,\omega} + E_n^{t,\omega} - O_n^{t,\omega} \qquad \forall a, \omega, t, n \qquad (4\text{-}9)$$

$$x_a^{t,\omega} = x_a^{t-1,\omega} + m_a^{t,\omega} - d_a^{t,\omega} \qquad \forall a, \omega, t \qquad (4\text{-}10)$$

$$x_a^{t,\omega} = \sum_{i}\sum_{j}\sum_{\tau \le t}(\hat{d}_{i,j}^{\tau}(-) * h_{i,j,a}^{\tau,t,\omega}) \qquad \forall a, \omega, t \qquad (4\text{-}11)$$

$$h_{i,i_v^{\omega},j}^{\tau,t,\omega} = \psi_2(f_{\omega}(\hat{d}_{i,j}^{\tau}(-)), \qquad \forall i, i_v^{\omega}, j, \tau t, t, \hat{d}_{i,j}^{\tau}(-), \omega \qquad (4\text{-}12)$$

$$\sum_{j}\sum_{k(v)} \hat{r}_{i_v^{\omega},j,k(v)}^{\tau,v,\omega} = \sum_{\tau \le t}\sum_{i}\sum_{j}(\hat{d}_{i,j}^{\tau}(-) * h_{i,i_v^{\omega},j}^{\tau,t,\omega}) \qquad \forall i_v^{\omega}, i, \omega, \tau \qquad (4\text{-}13)$$

$$\sum_{k(u)} T_{i,j,k(u)}^{\tau,u,\omega} = \sum_{\tau \le t}\sum_{a}\left(h_{i,j,a}^{\tau,t,\omega} * \Delta\right) \qquad \forall i, j, \omega, \tau, t, k(u) \qquad (4\text{-}14)$$

$$d_a^{t,\omega} = \sum_{i}\sum_{j}\sum_{\tau}\sum_{u}\sum_{k(u)} d_{i,j,k(u),a}^{\tau,u,t,\omega} \qquad \forall a, t \qquad (4\text{-}15)$$

$$m_a^{t,\omega} = \sum_{i}\sum_{j}\sum_{\tau}\sum_{u}\sum_{k(u)} m_{i,j,k(u),a}^{\tau,u,t,\omega} \qquad \forall a, t \qquad (4\text{-}16)$$

$$I_n^{\tau,\omega} = \sum_{j}\sum_{u} r_{n,j}^{\tau,u} \qquad \forall \tau, n \qquad (4\text{-}17)$$

$$O_n^{t,\omega} = \sum_{i}\sum_{c}\sum_{\tau}\sum_{u}\sum_{k(u)} m_{i,n,k(u),c}^{t,u,t,\omega} \qquad \forall t, c \in C(n), n \in J \qquad (4\text{-}18)$$

$$\tau \le t \qquad (4\text{-}19)$$

$$h_{i,i_v^{\omega},j}^{\tau,t,\omega} = 0 \; or \; 1 \qquad (4\text{-}20)$$

$$z_a = 0 \; or \; 1 \qquad (4\text{-}21)$$

$$All \; \mathrm{var}iables \ge 0 \qquad (4\text{-}22)$$

The objective function (4-1) is to maximize the long run average of the second-stage random values under stochastically occurred incident in the network. A recourse decision can be made in the second stage to correct the locations due to the change of traffic pattern caused by the random incidents. Note that the deterministic model is a special case of the proposed stochastic model without considering incident scenarios ($S = 0$).

Constraint (4-2) shows that the summation of weights of all objectives should equal to 1. Constraint (4-3) ensures that the total number of network sensors is within the budget/resource limitation.

Constraint (4-4) is the information gain contributed by link a through the observation of O-D pair w_1 during observation interval t under scenario ω. It denotes that the time-dependent link information gain is a function of time-dependent link proportion values (assignment matrix). The scenario-dependent link proportion value leads to a random recourse function. There are several details to note about the link information gain matrix $k_{w_1,a}^{t,\omega}$. First, it is the product of $P_{w_1,w_2}^{t} * h_{w_2,a}^{\tau,t,\omega}$. The a *priori* variance covariance matrix (P_{w_1,w_2}^{t}) indicates the existing estimated demand uncertainty level of each OD pair and covariance between different OD pairs. Assignment matrix ($\mathbf{H}$) connects link observations to the OD demand. If a link can intercept those OD pairs with a large variance, or a link can intercept more than one OD pairs, then the product of $P_{w_1,w_2}^{t} * h_{w_2,a}^{\tau,t,\omega}$ is more likely to be large and more demand uncertainty reductions are obtainable. The second detail about the link information gain matrix $k_{w_1,a}^{t,\omega}$ is the inverse of $h_{w_1,a}^{\tau,t,\omega} * P_{w_1,w_2}^{t} * h_{w_2,a}^{\tau,t,\omega}$. If one only considers $h^{\tau,t,\omega}_{w_1,a} * h^{\tau,t,\omega}_{w_2,a}$ for multiple possible sensors, the inverse of $h^{\tau,t,\omega}_{w_1,a} * h^{\tau,t,\omega}_{w_2,a}$ specifies the correlation of measurements among multiple links. Furthermore, $h^{\tau,t,\omega}_{w_1,a} * P^{t}_{w_1,w_2} * h^{\tau,t,\omega}_{w_2,a}$ describes the measurement correlation on the basis of existing estimated variance and covariance. If $h^{\tau,t,\omega}_{w_1,a} * P^{t}_{w_1,w_2} * h^{\tau,t,\omega}_{w_2,a}$ is large, meaning that either new senor data could be highly correlated with each other or they are correlated with the current estimate, then the inverse is small and the weight factor becomes insignificant. Recall that generalized linear regression (GLS) has the same term $(\mathbf{HPH}^T)^{-1}$, indicating the extent of information/knowledge obtained from observations.

Constraint (4-5) expresses the link proportion values as a function of network time-dependent link flows. Function $\psi_1(f_\omega(\hat{d}^{\tau}_{i,j}(-))$ is a complicated non-linear function, which embeds the impact of traffic link flow, routing policy, signal control, traffic demand, etc. on the link proportion values over a planning horizon. Analytically, the assignment matrix is determined by the route choice fraction and network traffic flow propagation (Cascetta *et. al* (1993)).

$$h^{\tau,t,\omega}_{w,a} = \sum_{u} \sum_{k(u)} (\alpha^{\tau,t,u,\omega}_{w,a,k(u)} * q^{\tau,u,\omega}_{w,k(u)})$$

where $\alpha^{\tau,t,u,\omega}_{w,a,k(u)}$ is the link-path incidence fraction for OD pair w under scenario ω, $q^{\tau,u,\omega}_{w,k(u)}$ is the average fraction of choosing path $k(u)$ at departure time τ for OD pair w under scenario ω. Based on the assumption that the vehicles are uniformly distributed in a packet and travel times are observable, Cascetta *et. al* (1993) derived a relationship between the link path incidence and travel time. Due to its dynamic complexity, link proportion values in this study are obtained from DYNASMART-P simulation results, as they would be in practical applications.

Constraints (4-6) and (4-7) state the time-dependent user equilibrium principle. With the non-negative path flow conservation, these two constraints are the general first-order conditions for the dynamic user equilibrium. The paths connecting node i_v^{ω} where the impacted vehicles reroute under scenario ω to any destination during any departure time can be divided into two categories: those carrying flow, on which the travel time must be minimal; and those not carry flow, on which the travel time must be greater than or equal to the minimal travel time.

Constraint (4-8) specifies the reduction of link capacity due to the occurrence of a stochastic incident on this link. The probability of incident occurrence follows a pre-specified distribution (Poisson distribution in this research). Constraint (4-9) denotes the node flow conservation under scenario ω. Constraint (4-10) represents the link flow conservation. It shows that flows on a link during observation time interval t are determined by the inflow, outflow and vehicles on that link during last time interval $t-1$.

Constraint (4-11) expresses that the number of vehicles on a link during any time interval is determined by the demand and the corresponding link proportion value $h_{i,j,a}^{\tau,t,\omega}$.

Constraint (4-12) expresses the time-dependent node-path incident variable. Similar to constraint (4-5), it is a non-linear function of traffic demand and determined by the interaction of different components, such as link traffic flow, incident characteristics, signal data, etc. It is obtained from DYNASMART-P simulation in this study. Constraint (4-13) expresses the number of impacted vehicles rerouting to any other destination at any time interval, as determined by the OD demand and the node-path incidence variable under scenario ω.

Constraint (4-14) states that the total travel time for an OD pair is the summation of the travel time along each possible path of that particular OD pair under scenario ω. Instead of assuming that the vehicles leaving an origin during any departure interval act like a single user (discrete packet approach), this study assumes that those vehicles are continuously spread over the interval between the "head" and "tail" of the packet (continuous packet approach) (Cascetta and Cantarella, 1991) and thus $0 \le h_{i,j,a}^{\tau,t,\omega} \le 1$. This assumption is more realistic than the time-average link flow assumption which assigns the link-path incidence fraction either 1 or 0. In addition, this assumption reflects more closely the simulation package's philosophy.

Constraints (4-15) and (4-16) represent the number of vehicles leaving and entering a link. Constraints (4-17) and (4-18) represent the number of vehicles leaving and entering a node.

Constraint (4-19) expresses that the departure time is always less than or equal to the current observation time. Constraints (4-20) and (4-21) define two binary integer variables. Constraint (4-22) makes sure all variables are non-negative.

4.3.2.2 Random Incident Generation Model

An incident generation model based on the model proposed by Chiu, Huynh and Mahmassani (2001) is used in this study to generate network random incidents. It is assumed that (1) occurrence of n incidents on link a follows Poisson process with occurrence rate λ; (2) the occurrence rate λ is identical on all the links; (3) each link has some probability of having an incident on it; and (4) the incidents are independent of each other. Due to different link congestion levels at different time intervals, the incident occurrence probability is different from time interval to time interval and from location to location. Then the probability that *n* incidents occur on link *a* during time interval t is

$$P_a^t(x=n)=\frac{(\lambda L_a f_a(t))^n\cdot e^{-\lambda L_a f_a(t)}}{n!}$$

where $f_a(t)$ is the link flow at time interval t on link a. For simplicity, this study assumes that the incident probability is not time-dependent and has the following expression:

$$P_a(x=n)=\frac{(\lambda L_a \widetilde{f}_a)^n\cdot e^{-\lambda L_a \widetilde{f}_a}}{n!} \tag{4-23}$$

where

λ: occurrence rate per unit length and unit flow of the network

L_a: length of link *a*.

$\widetilde{f}_a$: total volume of link flow across the simulation horizon T, $\widetilde{f}_a=\sum_{t\in T} f_a(t)$

The probability of an incident occurring in the network is

$$P_{all}(x=1)=(\lambda\sum_{i\in A}(\widetilde{f}_i\cdot L_i))\cdot e^{-\lambda\sum_{i\in A}(\widetilde{f}_i\cdot L_i)} \tag{4-24}$$

According to Bayes' theorem, the conditional probability that an incident occurs on link *a*, given there is one incident occurred in the network is as follows,

$$P_a(x=1 \mid w=1) = \frac{P_a(x=1) \cdot P_{-a}(x=0)}{P_{all}(w=1)} = \frac{\widetilde{f}_a \cdot L_a}{\sum_{i \in A}(\widetilde{f}_i \cdot L_i)} \tag{4-25}$$

Eq.(4-25) says that the incident occurrence probability on link a is the ratio of the weighted lane-miles of link a to the total weighted lane-miles if one incident happens. Thus, the likelihood of an incident occuring on a link is proportional to the link length, number of lanes and congestion level.

Similarly, the probability of one incident occurring on link a and the other incident on link b when two incidents happens is given by:

$$\begin{aligned} P_{a,b}(x=1, y=1 \mid w=2) &= \frac{P_a(x=1) \cdot P_b(y=1) \cdot P_{-a,-b}(x=0)}{P_{all}(w=2)} \\ &= \frac{2*(\widetilde{f}_a \cdot L_a) \cdot (\widetilde{f}_b \cdot L_b)}{(\sum_{i \in A}(\widetilde{f}_i \cdot L_i))^2} \end{aligned} \tag{4-26}$$

Note that the probability of two incidents occurring on the same link is

$$P_a(x=2 \mid w=2) = \frac{2*(\widetilde{f}_a \cdot L_a)^2}{(\sum_{i \in A}(\widetilde{f}_i \cdot L_i))^2} \tag{4-27}$$

The above results show that links with longer length, more lanes and larger flow exhibit higher incident occurrence probability.

4.3.2.3 Deterministic equivalency of SOSLP model

$E_\xi(Q(\mathbf{H}(\omega, \mathbf{Z})))$ is the expected OD coverage and link information gain under different scenario ω, i.e. one incident, two incidents, three incidents, etc. Under finite discrete distribution assumption of the random scenarios, SOSLP can be formulated as a deterministic equivalent program as follows:

$$
\begin{aligned}
&J = Max\{E_{\xi}(Q(\mathbf{H}(\omega,\mathbf{Z})))\} \\
&s.t. \\
&\qquad Constraints\ (4-2) \sim (4-22)
\end{aligned}
\tag{4-28}
$$

where

$$
\begin{aligned}
E_{\xi}(Q(\mathbf{H}(\omega,\mathbf{Z}))) &= \sum_{\omega=0}^{S}\{P(\xi=\omega)(Q(\mathbf{H}(\omega,\mathbf{Z}))\mid\xi=\omega)\} \\
&= P(w=0)[(Q(\mathbf{H}(\omega,\mathbf{Z}))]+ \\
&\quad P(w=1)\left[\sum_{a\in A}P_a(x=1\mid w=1)(Q(\mathbf{H}(\omega,\mathbf{Z}))\mid\xi_a^t)\right]+ \\
&\quad P(w=2)\left[\sum_{a,b\in A}P_{a,b}(x=1,y=1\mid w=2)(Q(\mathbf{H}(\omega,\mathbf{Z}))\mid\xi_a^t,\xi_b^t)\right]+ \\
&\quad P(w=3)\left[\sum_{a,b,c\in A}P_{a,b,c}(x=1,y=1,z=1\mid w=3)(Q(\mathbf{H}(\omega,\mathbf{Z}))\mid\xi_a^t,\xi_b^t,\xi_c^t)\right]+ \\
&\quad \cdots
\end{aligned}
\tag{4-29}
$$

The above deterministic equivalent model converts the SOSLP to a mixed integer non-linear model. The integer L-Shaped based algorithm or local search heuristics, such as simulated annealing or Tabu search can be applied to this problem. Unfortunately, for a large scale network and its induced thousands of realizations, the L-shaped method would consume greater computational resources to solve the complicated linear problem and require additional attention to decomposition techniques, such as Benders' decomposition, to take advantage of the model structure. Within Benders' framework, two different types of linear programming models would need to be solved: a master problem that solves for the first stage variables, and a series of sub-problems that deal with second stage variables. Although the sub-problems in the SOSLP are always feasible, the SOSLP is not a convex problem due to the complicated dynamic characteristics of the assignment matrix. This provides the motivation to develop and test heuristics, which can find robust solutions to the given large-scale non-convex stochastic program.

4.4 Solution Procedure

The hybrid Greedy Randomized Adaptive Search Procedure (HGRASP) solution procedure proposed in the previous chapter to solve the deterministic model is modified to search for the best solution for the stochastic model. The candidate sensor locations are evaluated by the multiple user class procedure integrated in the DTA assignment simulation tool, DYNASMART-P (Peeta & Mahmassani 1995). Once an incident realization (ω) is detected, the affected vehicle paths and theis associated zones are delineated. All newly generated vehicles (during the incident) from these impacted origin zones and the en-route impacted vehicles that would have originally traversed the incident link are classified as user class v^{ω}, provided with diversion guidance to take such routes that minimize their travel time. All other vehicles will be classified as user class o^{ω} and will retain their original assigned paths. The next section illustrates the modified hybrid greedy randomized adaptive search procedure.

4.4.1 Hybrid Greedy Randomized Adaptive Search Procedure (HGRASP)

Greedy Randomized Adaptive Search Procedure is a multi-start or iterative sampling method (Lin & Kernighan, 1973, Feo & Resende 1995, Festa & Resende 2001, Pitsoulis and Resende 2001), with each GRASP iteration composed of two phases, a solution construction phase, where a randomized greedy solution is constructed, and a solution improvement (local search) phase, which starts at the constructed solution and applies iterative improvement until a locally optimal solution is found. Figure 4-2 depicts the solution procedure flow chart for the SOSLP, which is summarized in the following steps:

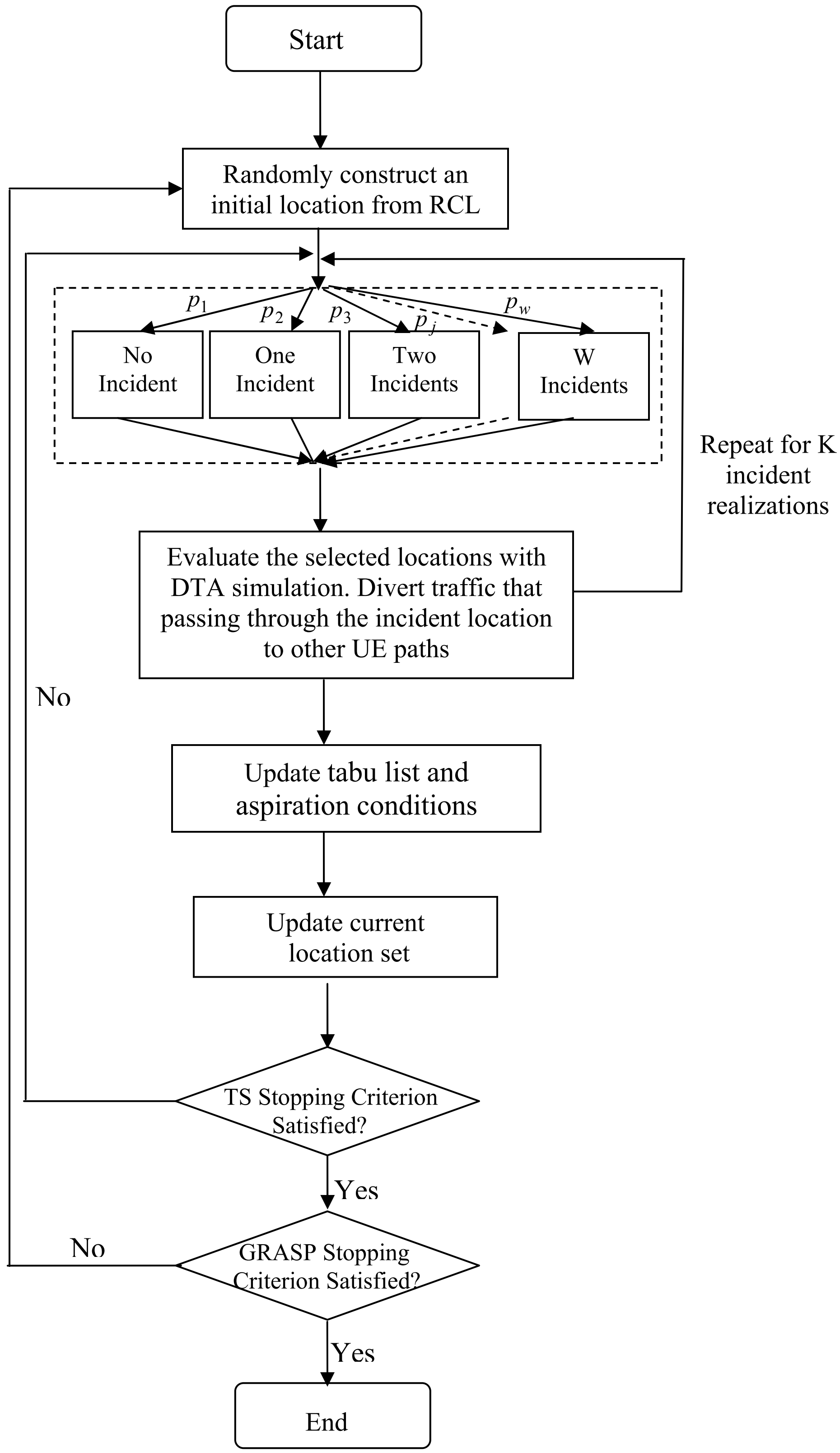


Figure 4-2 Hybrid GRASP-DTA bi-level solution procedure for SOSLP

HGRASP-DTA Solution Procedure for SOSLP

Step 0 (Initialization): Set $F^* = F(\mathbf{Z}^*) = -\infty$, where $\mathbf{Z}^*$ is the solution vector representing the best locations found so far.

Step 1 (Construction & Searching): Repeat if GRASP stopping criterion is not satisfied.

(a). Construct a greedy randomized solution $\mathbf{Z}$

(b).Local Search (Tabu Search): find local optimal vector $\mathbf{Z}'$ in the neighborhood $N(\mathbf{Z})$

1. Generate random incident realization

   1a).Draw a random number α from a uniform distribution (0,1) $\alpha \in UNIF[0,1]$, and map it to the corresponding Poisson distribution probability p_j to generate number of incidents scenario($w = j$);

   1b).Draw a random number β from a uniform distribution (0,1) $\beta \in UNIF[0,1]$, and map it to the corresponding conditional Poisson distribution probability $P_{a_1,.a_j}(a_1 = 1,\cdots,a_j = 1 \mid w = j)$ on links;

2. Evaluate the selected sensor locations $\mathbf{Z}'$ with DTA simulation. Divert traffic that passes through the incident location to other UE paths

3. Go back to 1 and repeat for k incident realizations

(c). Update tabu list and solution: if $F(\mathbf{Z}') > F^*$, *let* $F^* = F(\mathbf{Z}')$ *and* $\mathbf{Z}^* = \mathbf{Z}'$. If Tabu search stopping criterion is not satisfied, go to (b), otherwise go to (a).

Step 2 (Best Solution Found): Return the best locations found $\mathbf{Z}^*$

In the construction phase, the candidate locations, ranked with respect to a greedy function which measures the benefit of choosing each location, are randomly selected one by one at each time. Pitsoulis and Resende (2001) summarized different random element selection methods to build a list of best candidates but not necessarily the top candidates during every HGRASP iteration. The list is called a restricted candidate list (RCL). This selection technique enables the heuristic to diversify the exploration in the search space. In this study, a randomly generated $\alpha \in UNIF[0,1]$ value coupled with an adaptive greedy function were used to build the RCL at each HGRASP iteration. Below is the procedure followed in the construction phase:.

Construct a greedy randomized solution $\mathbf{Z}$

Step 0 (Initialization): *Set* $\mathbf{Z} = \{ \}$

Step 1 (Construction): Repeat until the total elements in set $\mathbf{Z}$ equal to the number of sensors L

(a). $\hat{c}_{\max} = Max\left\{ \hat{c}^l(t) \mid \hat{c}^l(t) = \sum_{\tau \le t} \sum_{w \in W} (h_{w,l}^{\tau,t,\omega_0} * \hat{d}_w^{\tau}(-)), l \in A \right\}$, where $\hat{c}_{\max}$ is the maximal link flow under the normal traffic condition across the entire planning horizon T

(b). $RCL = \{ l \in A \mid \hat{c}^l \ge \rho * \hat{c}_{\max} \}$, where $\rho \in [0,1]$ is a scalar.

(c). Pick l at random from RCL, while $l \notin \{ L_k \mid L_k \in R_w^t(l_z), \forall l_z \in Z, t \le T, w \in W \}$

(d). $\mathbf{Z} = \mathbf{Z} \cup \{l\}, A = A \setminus \{l\}$

Step 2: Return the solution set $\mathbf{Z}$

$R_w^t(l_z)$ is a set of paths that traverse link l_z and connect O-D pair w during time interval t; L_k denotes the set of links on path k. Step 1(c) shows that the candidate link l cannot be on any path k that traversed those selected links in set $\mathbf{Z}$. The inherent idea in step 1(c) is to select links that can contribute greater information gains while keeping the rank of assignment matrix $\mathbf{H}$ full. By keeping the selected links uncorrelated, the procedure can obtain more information, as described in Section 3.4 in conjunction with the DOSLP.

Again as with the DOSLP, the solutions from the HGRASP-DTA construction phase are usually not locally optimal, and a local search procedure is employed to exploit the neighborhood $N(\mathbf{Z})$ of solution $\mathbf{Z}$ during every HGRASP iteration. A similar Tabu search procedure is applied here as well. The steps are repeated for completeness:

Local Search: finding local optimal vector $\mathbf{Z}'$ in the neighborhood $N(\mathbf{Z})$

Step 0 (Initialization): *Set* $k = 0$, empty the tabu list

Step 1: Repeat until the stopping criterion is satisfied

(a) (*Drop Move*). Randomly choose a location $x \in \mathbf{Z}$

(b) (*Add Move*). Set $k = k + 1$, $N(x,k)$ is the path set of the neighborhood of x at step k, where $N(x,k) = \{l \mid l \in R_{(i,j)}^t(x), 0 \le t \le T\}$.

A logit formulation is used to determine the selection probability, which let all of the links likely be selected while those links with larger flows have higher likelihood to be selected. Therefore, any link with flow has the probability to be selected, but those links with higher congestion were more likely to be selected.

$$P_l = \frac{e^{\alpha \cdot \hat{c}^l}}{\sum_{i \in N(x,k)} e^{\alpha \cdot \hat{c}^i}} \qquad (4\text{-}30)$$

where P_l is the probability of choosing link l

$\hat{c}^l = \sum_{t \le T} \sum_{\tau \le t} \sum_{w \in W} (h_{w,l}^{\tau,t,\omega_0} * \hat{d}_w^{\tau}(-)), l \in N(x,k)$ is the summation of simulated link flows on link l in planning horizon T

α is a scaling parameter

Scanning the tabu list, if the selected link l is not on the list or if the selected link l is on the list, but aspiration criteria is satisfied, put this link l at the bottom of the list. Otherwise, ignore this link and choose another link l', Set $\mathbf{Z}' = (\mathbf{Z} / \{x\}) \cup \{l\}$

(c) (*Update*). If $F(\mathbf{Z}') > F(\mathbf{Z})$, *Set* $\mathbf{Z} = \mathbf{Z}'$, $F(\mathbf{Z}) = F(\mathbf{Z}')$, update the tabu list and aspiration conditions.

(d) If $k \le K_{tabu}$, where K_{tabu} is the maximal tabu iterations, goto (a), otherwise, go to step 2

Step 2: Return the local optimal solution set $\mathbf{Z}$

The proposed HGRASP-DTA heuristic starts from a set of initially uncorrelated locations that intercept the largest OD flows, and iteratively explores the neighborhood of current solution till the stopping criteria being satisfied. However, the decision makers' preference to reduce system uncertainty or increase O-D flow coverage in the long run affects the final sensor placement.

4.5 Summary

Uncertainty is pervasive in transportation planning and has a significant influence in the transportation evaluation and decision making. With particular emphasis on the time-dependent OD demand estimation problem under a variety of a *priori* unknown incident scenarios, this chapter proposed a two-stage stochastic model with recourse to find an optimal set of sensor locations, subject to a budget constraint, with the dual aim of maximizing the long run expectation of the link information gains and the OD flow coverage in a large scale traffic network The proposed model is based on the time-dependent link measurement equations, with the aim of minimizing the deviation between the simulated and observed link counts by considering different error sources, such as link measurement errors, estimation errors, and etc. A modified HGRASP-DTA search procedure is used to find the near optimal sensor locations in the context of dynamic traffic assignment and stochastic scenarios.

Chapter 5 Sensitivity and Experimental Analysis of Sensor Location Problem Methods

5.1 Introduction

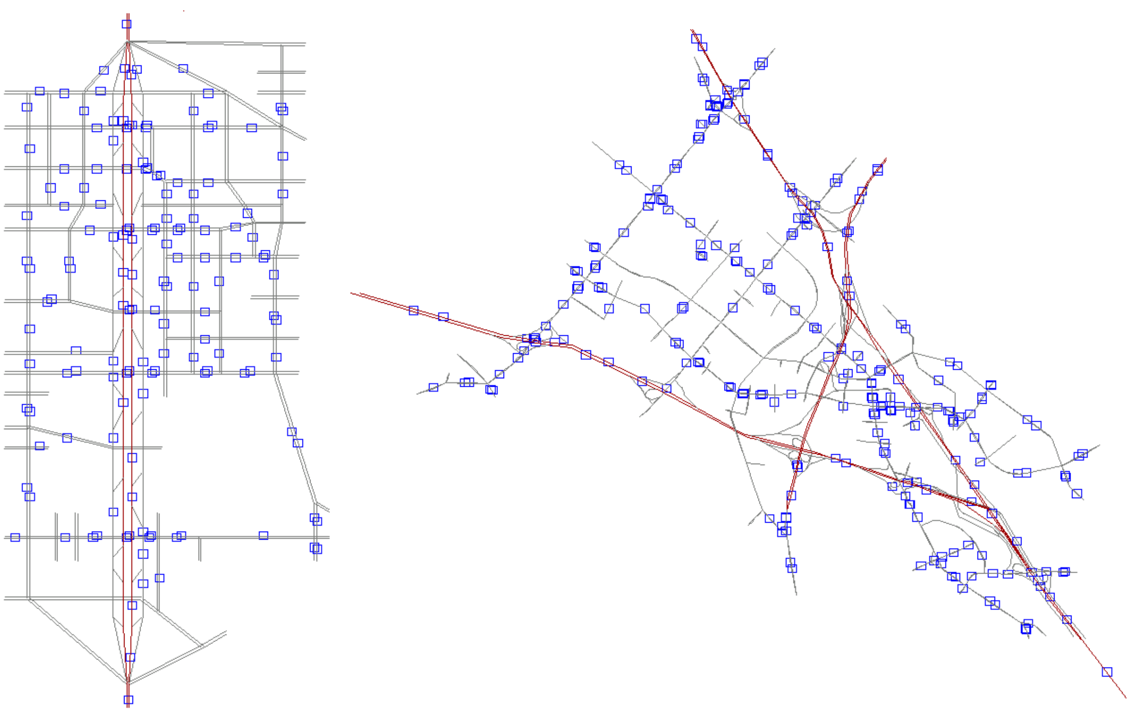
This chapter aims to evaluate the performance of the proposed models under different conditions in terms of the value of available information from deployment of network sensor locations. With regards to the complexity of the assignment matrices in the context of real-time traffic estimation and prediction, simulated assignment matrices that obtained from a dynamic traffic assignment based simulation software can circumvent the complexity of the analytical derivation and are used in the proposed models in order to capture the network traffic patterns and dynamics. The sensitivity analysis of estimation and prediction quality is conducted in this chapter using the DYNASMART-X real-time DTA system. The analysis considers both randomly generated location scenarios as well as scenarios based on engineering judgment. The latter considers placing sensors on high volume links on the main freeways and arterials. Taken together, the two sets of scenarios provide useful insight into the robustness of the real-time DTA estimation and prediction, and the effect of location-specific considerations on estimation and prediction quality. The DOSLP and SOSLP models are tested on an actual large-scale network. The results are evaluated and compared with those from the sensitivity studies to assess the respective performance of the proposed models. The value of additional information from a new sensor in traffic status estimation and prediction is also characterized in terms of its contribution to the demand uncertainty reduction.

The principal objectives of this chapter include: (1) illustrate the effectiveness of the proposed models with the real-world application using actual data; (2) evaluate the optimal sensor locations derived from the static model and dynamic model; (3) determine the marginal value obtained from each additional sensor, in terms of the demand estimation errors and OD flow coverage in the context of traffic dynamics; (4) demonstrate the influence of the network uncertainty on the sensor locations; and (5) perform sensitivity analyses to assess the robustness of the estimated demand matrix to the sensor numbers and locations. The proposed methodologies are expected to provide insight on optimal sensor deployment in large scale networks for real-time traffic estimation and prediction.

This chapter is organized as follows. Section 5.2 evaluates the performance of the proposed set covering model under the assumption of an unlimited number of sensors for two medium-size networks, and thereafter scenarios under a limited number of sensors are tested. Section 5.3 evaluates the performance of optimal sensor locations derived from different methods with budgetary constraints in a large scale network. It starts with sensitivity studies with respect to the number and location of the sensors in terms of impact on the traffic estimation and prediction under real-time information. Next, the results obtained from the proposed DOSLP and SOSLP methodologies are analyzed under stochastic and deterministic scenarios. Finally, the major conclusions are summarized.

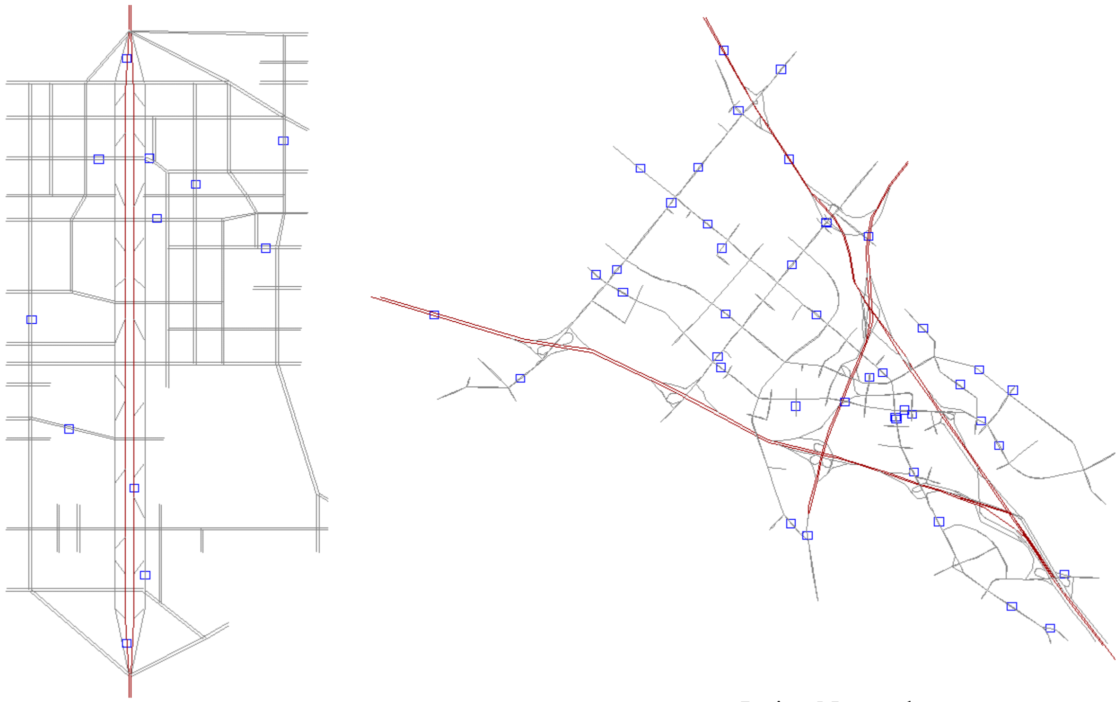
5.2 Unlimited Network Sensors for two Medium-Size Networks

In order to illustrate the proposed OD covering model, Figure 5-1 shows the sensor locations for two networks: 1) Fort-Worth, TX, with 147 sensors that cover 156 OD pairs (13 TAZ), including 180 nodes and 445 links and 2) Irvine, CA, with 238 sensors that cover 3660 OD pairs (61 TAZ) , including 326 nodes and 626 links. The *a priori* "*relevant degree*" ζ^{τ} =0 under the dynamic traffic assignment. The time period of interest is the morning peak from 6:30AM-8:30AM. Figure 5-2 shows the solution results for the static model proposed by Yang *et al.* (1998). The same networks using static information result in having 12 sensors and 44 sensors respectively.



Fort-Worth   Irvine Network

Figure 5-1 Sensor Locations by DTA in Fort-Worth & Irvine Network



Fort-Worth Irvine Network

Figure 5-2 Sensor Locations by Static Model in Fort-Worth & Irvine Network

The results of the dynamic model show that due to the traffic dynamics, more sensors are needed in order to cover each OD pair in the network across time than those obtained by solving the sensor location problem based on static traffic assignment. Figure 5-3 shows the minimum number of required sensors for each departure time interval τ over the analysis horizon.

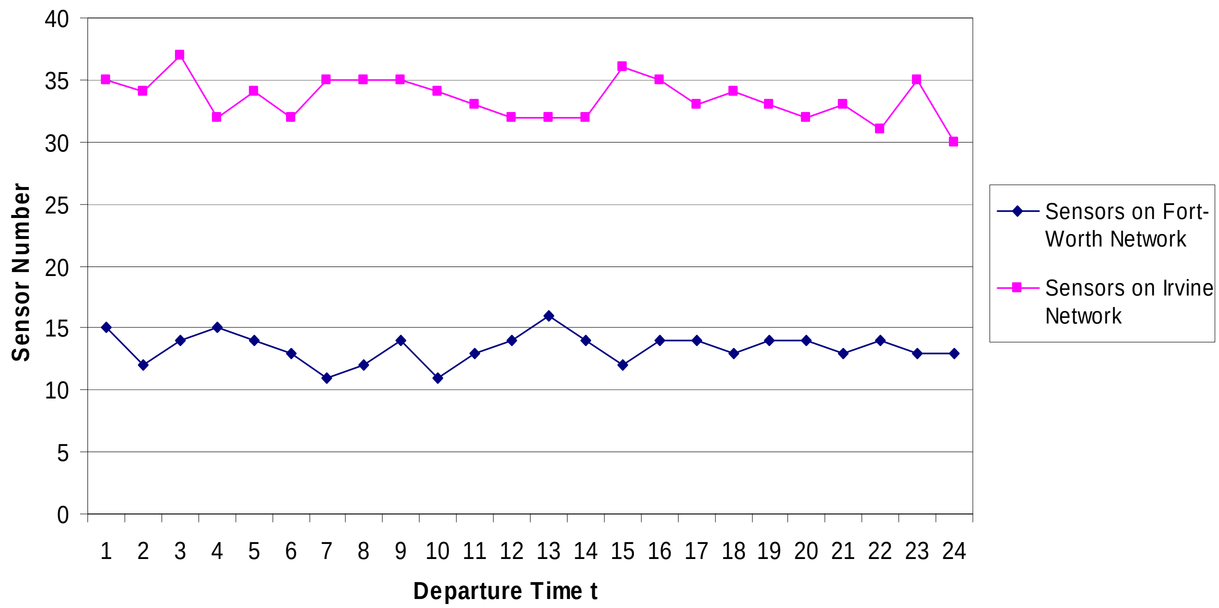


Figure 5-3 Number of Sensors for Each Time Period

5.2.1 Sensitivity Analysis on the Number of Sensors and Percentage OD Coverage

A sensitivity analysis is performed to explore the relationship between the number of sensors and level of OD coverage in a network. The purpose of this analysis is to explore the marginal value, in terms of percentage coverage, of adding sensors to the network. The analysis also provided a platform to investigate the effect of sensor location on the OD demand coverage rate.

By setting an appropriate ζ^{τ} in each departure time interval τ and solving the corresponding DOSLP -1 model, Figure 5-4 shows the different sensor numbers required to provide different levels of OD coverage in the Fort-Worth, TX and Irvine, CA networks under the dynamic model. As expected, to cover more OD pairs, more sensors have to be installed in the network. These results also indicate that obtaining greater than 50% OD coverage for either network require a significant increase in the number of sensors. In addition, the results show that a fairly low number of judiciously-placed sensors can provide a substantial amount of coverage.

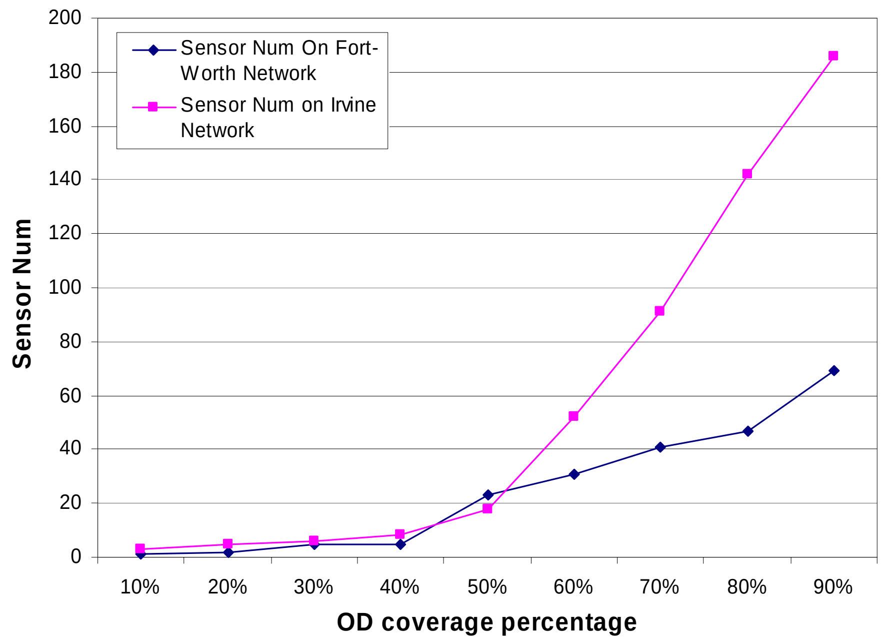
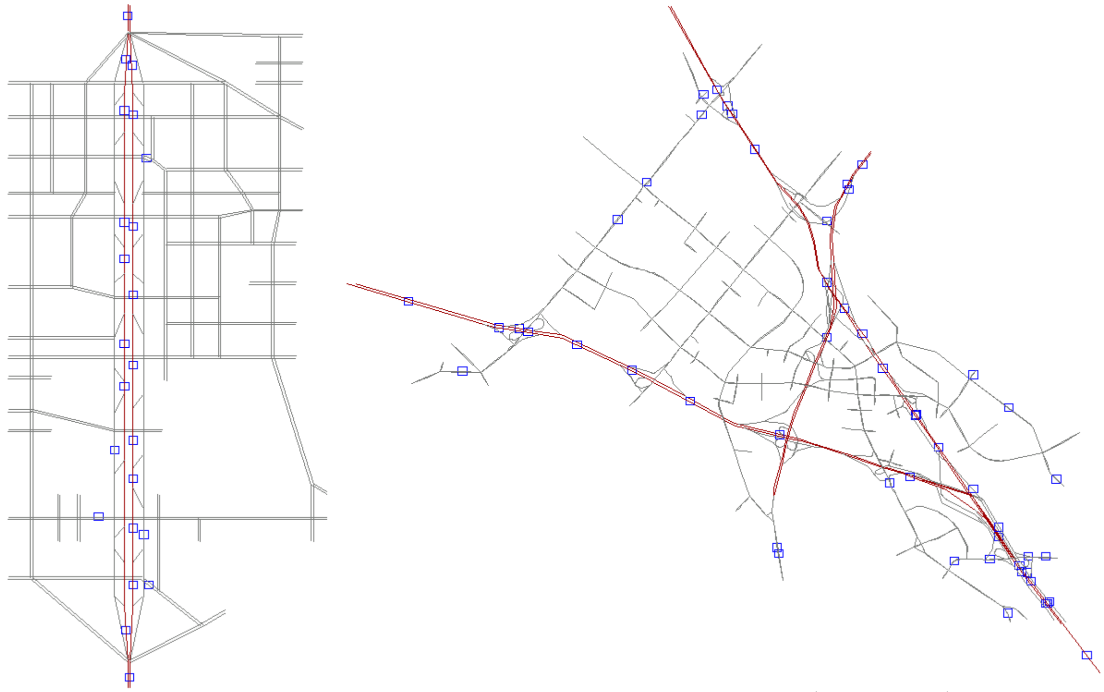


Figure 5-4 Sensors Covering Percentage OD Demand

Figure 5-5 shows 23 sensors covering 50% of the O-D demand flow on the Fort-Worth, TX test network, and 52 Sensors covering 60% of the O-D demand flow on the Irvine, CA test bed network. Interestingly, the sensors are mostly distributed along the freeways, in which the links have higher flows than that on the arterial streets. The results reveal that if budget is constrained, deploying sensors along the freeway would make sense in terms of maximization of the O-D demand coverage.



Fort-Worth

Irvine Network

Figure 5-5 Partial OD Demand Coverage on Different Network

5.3 Limited Network Sensors for a Large-Scale Network

This section evaluates the methodologies under the assumption of a limited number of sensors. First the sensitivity analyses of estimation and prediction quality vis a vis both sensor location and sensor coverage percentage in a network are performed, and then the solutions from the deterministic model (DOSLP) and stochastic model (SOSLP) are analyzed in stochastic and deterministic scenarios respectively. To illustrate the effects of network uncertainty on the sensor locations, the sensor locations and network performance from the deterministic model (DOSLP) are compared to those obtained by solving the sensor location problem based on the stochastic model (SOSLP) under different scenarios.

5.3.1 Maryland CHART Network Description

The experiments are performed on the CHART network in Maryland which was developed for use in real-time traffic management. Started in the mid 80's, CHART (Coordinated Highways Action Response Team) is the highway incident management program of the Maryland State Highway Administration (MDSHA). The study area is concentrated on the area surrounding the I-95 corridor between Washington, D.C. and Baltimore, MD. The network is bounded by I-695 to the north, I-495 in the south, US 29 in the west and I-295 in the east. The network includes four main freeways (I-95, I-295, I-495 and I-695), as well as two main arterials (US29 and Route 1). The Maryland CHART network reduces to 2,182 nodes, 3,387 links and 111 zones. It also includes 262 signals. Figure 5-6 shows the Maryland CHART network and signal locations. There are 14 working loop detectors deployed in the CHART study area. The locations of these detectors are shown in Figure 5-7. Ten of the detectors are located on I-95, two are located on I-495 and another two are located on MD-32. The detector information is frequently invoked in processing and interpreting the real-time traffic data and the actuated signal data. The time horizon of interest is the morning peak from 6:30AM to 8:30AM during which there are totally 119,189 vehicles generated. The DYNASMART-P simulation-based traffic assignment tool (Mahmassani *et. al* 2000) is used to load the time-dependent OD demand onto the network and assign paths to the vehicles.

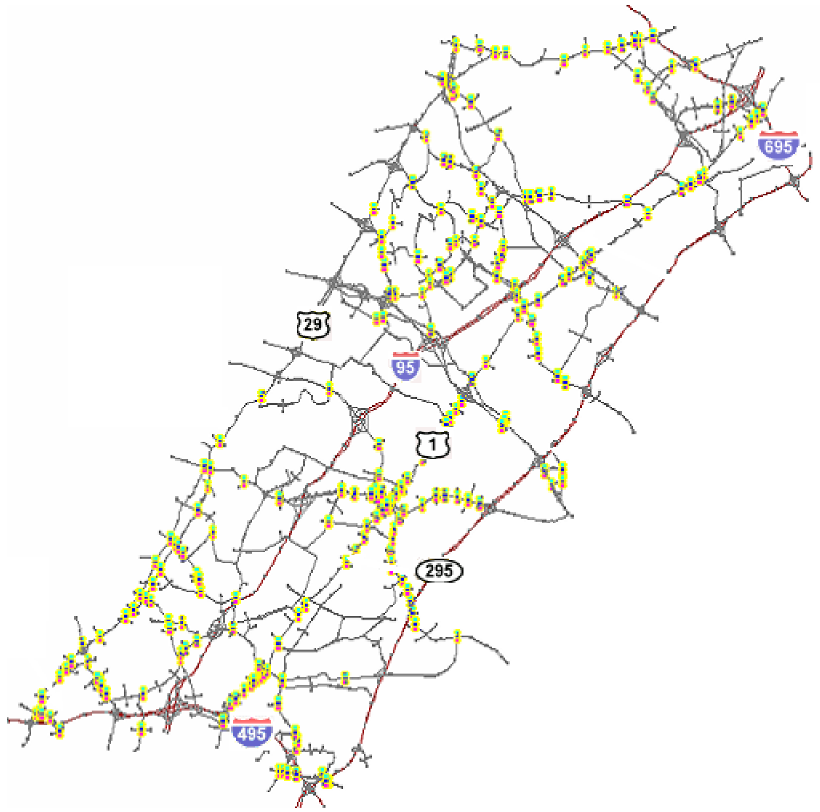


Figure 5-6 Maryland CHART Network

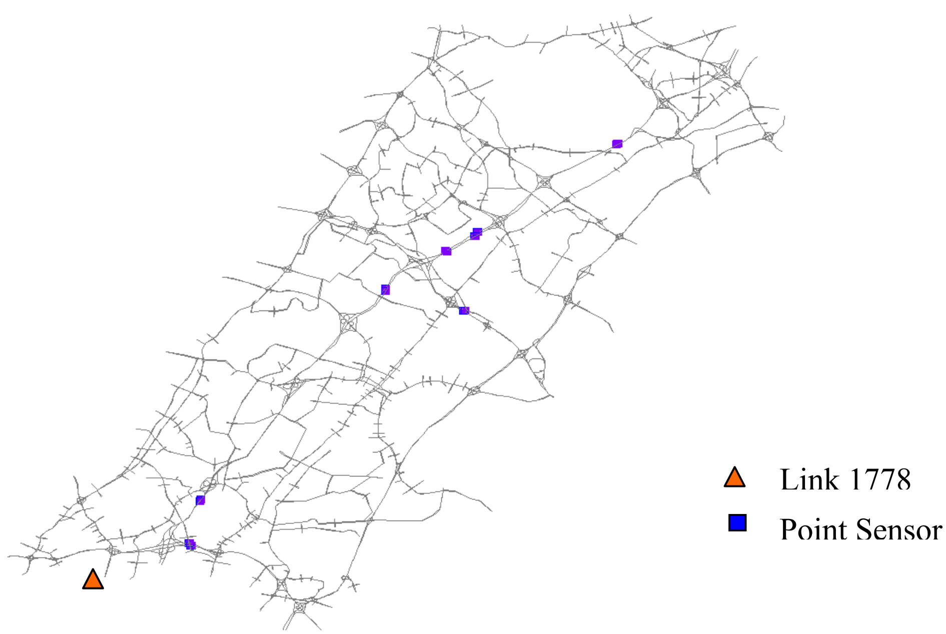


Figure 5-7 Existing Sensor Locations in Maryland CHART Network

5.3.2 Sensitivity Analysis of the Sensor Location and Estimated OD Matrix Quality

Two types of sensitivity analysis of estimation and prediction quality vis a vis both sensor location and level of sensor coverage in a network were conducted in this section. First a number of random sensor location scenarios were generated and analyzed. This set of analyses illustrates how the number of sensors in the network can influence the estimation and prediction results and also how distribution in the network can produce various results. A second set of location scenarios were generated using engineering judgment to place sensors on high volume links focusing on the main freeways and arterials. The analysis is conducted using the simulation assignment based DYNASMART-X real-time traffic simulator on the Maryland CHART network. The purpose is to explore the significance of adding sensors to the network and also to learn how their location affects estimation and prediction performance.

There are three parts in this section. The first part explains the procedure used to construct the sensor information, when observation data was not available. The second part introduces an analysis measure on the sensor locations and numbers to the link performance estimation and prediction, and the third part presents scenario descriptions and results from different scenarios are analyzed.

5.3.2.1 Experiment Data Synthesis

Within the study there are only 14 existing loop detectors (figure 5-7). These detectors collect and report data in 5-minute intervals. This detector information can be obtained from Center for Advanced Transportation Technology (CATT), Maryland DOT and Maryland SHA. Information describing detector location, as well as detector data is available from the CATT laboratory webpage (CATT 2004).

Each detector data file contains timestamp information, detector location, traffic direction, vehicle counts, vehicles/hour, speeds, and percent occupancy. Sensors collect 24-hour data in 5 minute intervals. The percent occupancy refers to the percentage of time the detector was occupied during the 5 minute interval. The speed is the average speed recorded over the 5 minute interval. The vehicles/hour is the 5 minute vehicle count converted to an hourly flow rate (ex. count =120, vehicles/hour = (120 vehicles/5 minutes)*(60 minutes/hour) = 1440 vehicles/hour). The vehicle count is the number of vehicles observed during the 5 minute interval.

The DYNASMART-X prototype is calibrated and evaluated according to its overall system functionality, rather than its individual modules, using the available data, with possible enrichment from other sources. The primary areas of calibration/evaluation are traffic estimation, traffic prediction, consistency checking/updating, and OD estimation/ prediction. Calibration and evaluation are performed at the overall system level. Calibration itself is separated into two types: a *priori* calibration of structural relations, and real-time adaptive updating of the calibrated models and parameter values. For these purposes a set of real-time data pre-processed was developed from the CATT laboratory databases. Necessary checking and judgment were exercised to retain consistency between the raw data and the pre-processed data. Data for the 14 links with reliable real-time data were processed for October 28 and November 1 - 5, 2004. This data was used as the basis of the network calibration and validation.

For the experiments conducted in this research, limited real-time data were available. Therefore experimental data that is used to mimic real-time sensor information was synthesized using a dynamic traffic assignment methodology (i.e. DYNASMART-P). To start, there is a time-dependent OD demand table, estimated using link counts coupled with a historical static demand table. This matrix is treated as the "ground truth" for experimental purposes. The ground truth OD demand is loaded onto the network using a dynamic traffic assignment simulation program to generate both link counts and density (simulated link measurements). The values become the "sensor data" or "observations" in the synthetic data set.

Note that to ensure the internal consistency between link flow measurements and density measurements, this study uses simulated link measurements as estimation input, instead of the actual link observations from the field data.

5.3.2.2 Analysis Measures

In order to interpret the influence that a given set of sensors has on the ability to estimate and predict network flow patterns, the root mean squared error (RMSE) of the link densities will be calculated for "all" of the links in the network. Note that generation links will not be included in these calculations. The calculation is as follows:

$$RMSE = \sqrt{\frac{\sum_{l,t}\left(C_{l,t} - C'_{l,t}\right)^2}{|L| \times T}}$$

where,

$C_{l,t}$ = observed density for link l during time interval t (ground truth output)

$C'_{,l,t}$ = simulated density for link l during time interval t (simulated output)

L is a set of links used in statistical calculations; $|L|$ is the total number of links in the set

T is number of time intervals

In a given scenario, the RMSE is calculation across all the links and across all of the time intervals.

5.3.2.3 Sensor Analysis Results

Each set of experiments was performed using a 6-hour simulation from 4AM to 10AM. To ensure that network loading and network discharge do not unduly influence the results, the analysis period was reduced to 5 hours (4:30AM to 9:30AM).

In developing the sensor location scenarios a few constraints were placed on the selection process. First the links sorted based on flow and the links with higher flow were considered to be more attractive. In addition, when consecutive links do not have access points between them, only one of the links was selected. Also, if a link is selected for sensor location, the two upstream and downstream links were not selected for sensor placement. The two rules were implemented in order to reduce correlation in the selected links and produce larger coverage of the network. The adjacency rules were not applied to ramps that connected to freeway links which had sensors. These selection constraints will be referred to as "filters".

5.3.2.3.1 Random Sensor Location Analysis

The first set of experiments is focused on 20 scenarios in which sensors are placed in the network on the basis of "random" selection. The selection of the sensor locations was not entirely random, in that they were selected at random from a subset of filtered links. This subset included the top 220 links sorted by link flow and filtered to meet the selection constraints. These 20 scenarios are described below:

Scenario 1-5: 5 runs with 20 sensors chosen randomly from the top 220 filtered sensors

Scenario 6-10: 5 runs with 30 sensors chosen randomly from the top 220 filtered sensors

Scenario 11-15: 5 runs with 40 sensors chosen randomly from the top 220 filtered sensors

Scenario 16-20: 5 runs with 80 sensors chosen randomly from the top 220 filtered sensors

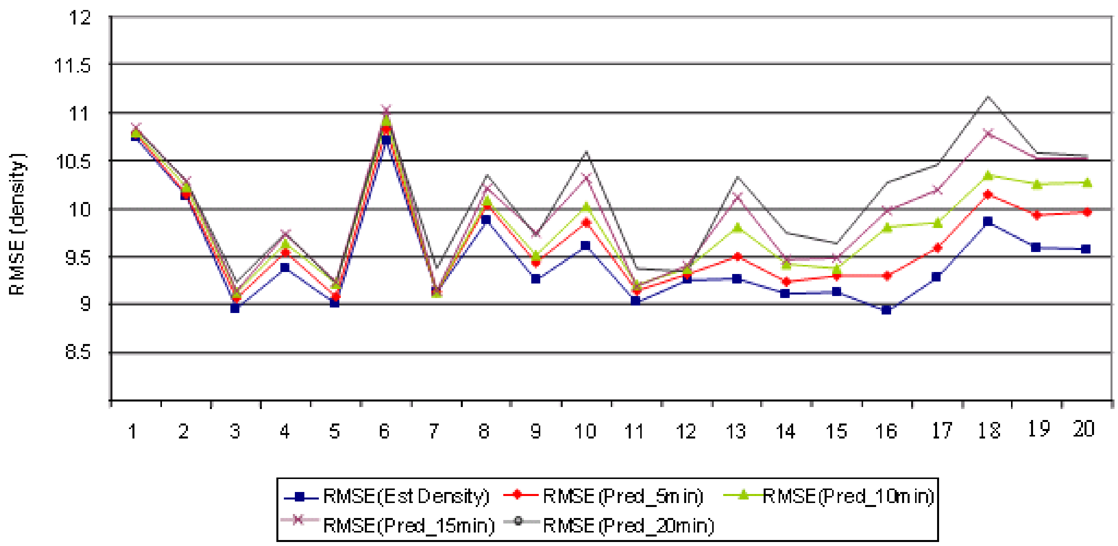


Figure 5-8 RMSE for Randomly Selected Sensor Locations

Figure 5-8 shows a plot of the RMSE for the estimation and prediction for each of the scenarios. From the figure one can observe the effects of the random location selection. Within each level of detection (i.e. 20, 30, 40 and 80 sensors) the random locations clearly produce variations in the results. In terms of estimation, Scenario 1 (20 sensors) is performing the worst, followed by Scenario 6 (30 sensors) and Scenario 2 (20 sensors). Also in terms of estimation Scenario 16 (80 sensors) is performing the best, followed by Scenario 3 (20 sensors) and Scenario 5 (20 sensors). Figures 5-9 -5-14 depicts the locations of these sensors in the network. The fact that two of top three *best* and *worst* scenarios in this analysis have with 20 sensors, emphasizes the value of good sensors placement. Given the ability to place 20 sensors in the network one would aim to place them to achieve the best results and not misplace them and obtain the worst.

In the case of the three worst scenarios (Figures 5-9--5-11), each of these scenarios lacks significant coverage on I-95 (the freeway with the most traffic).

Scenario 2 is performing the best out of the three and has the most coverage on I-95, as well as 10 additional detectors. Examining the three best scenarios (Figures 5-12-- 5-14), there are also a few commonalities. The most obvious is that each of these scenarios provides significant detection on I-495 (the east/west freeway at the southern edge of the network). In addition, each of these scenarios appears to provide detection at or around freeway access points throughout the network.

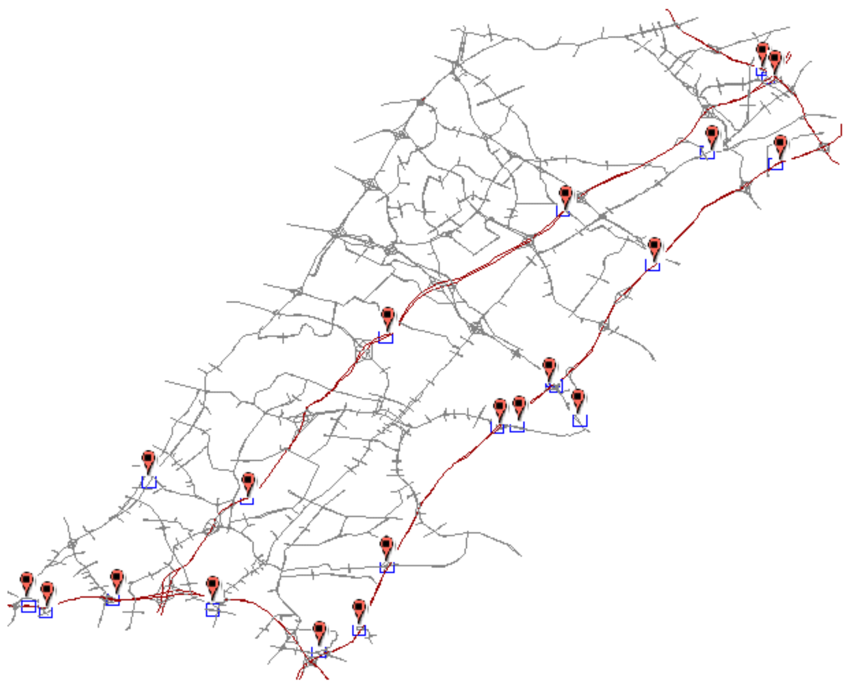


Figure 5-9 Scenario 1 (20 Sensors) Sensor Locations

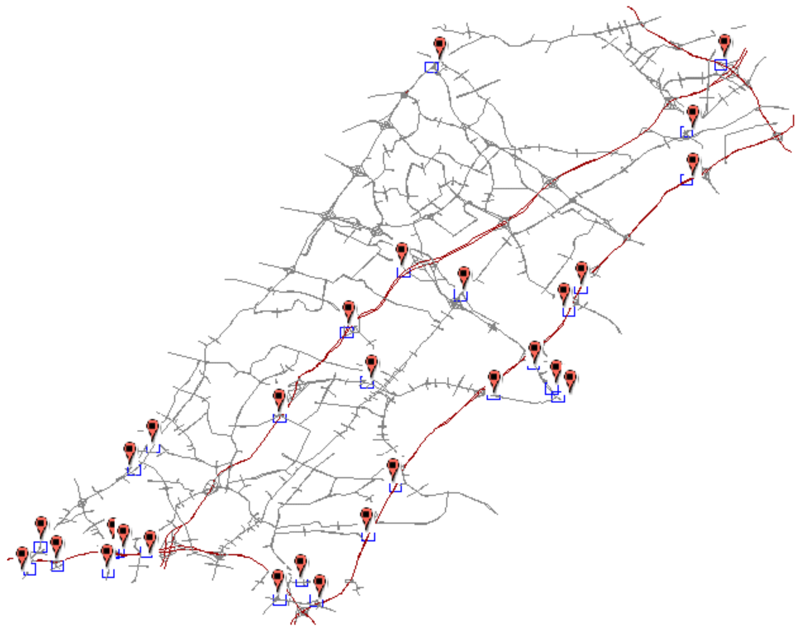


Figure 5-10 Scenario 6 (30 Sensors) Sensor Locations

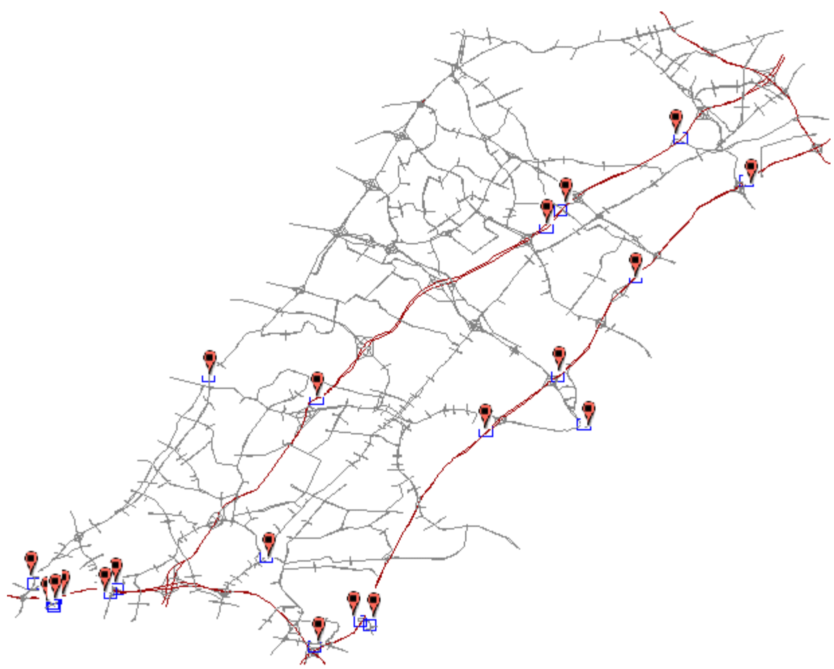


Figure 5-11 Scenario 2 (20 Sensors) Sensor Locations

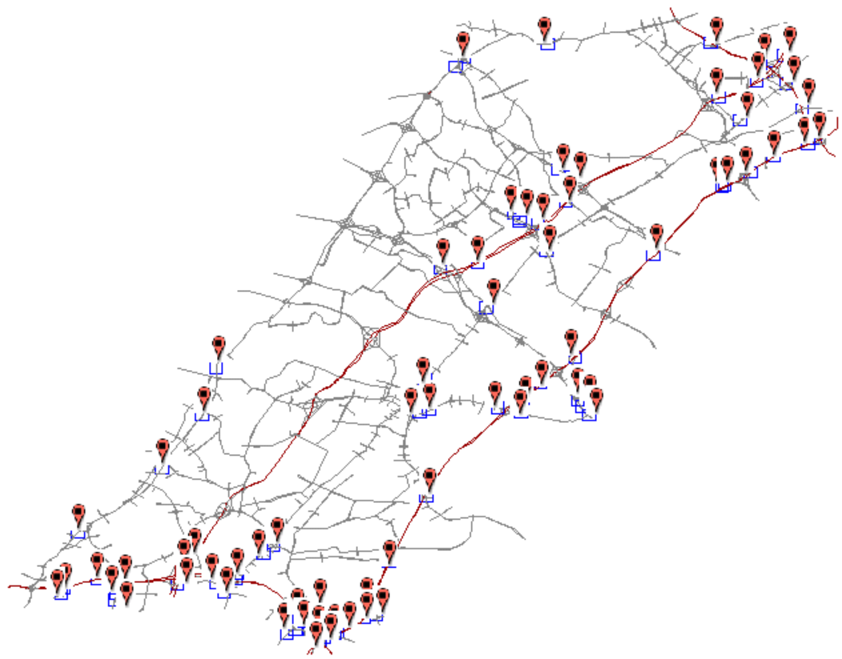


Figure 5-12 Scenario 16 (80 Sensors) Sensor Locations

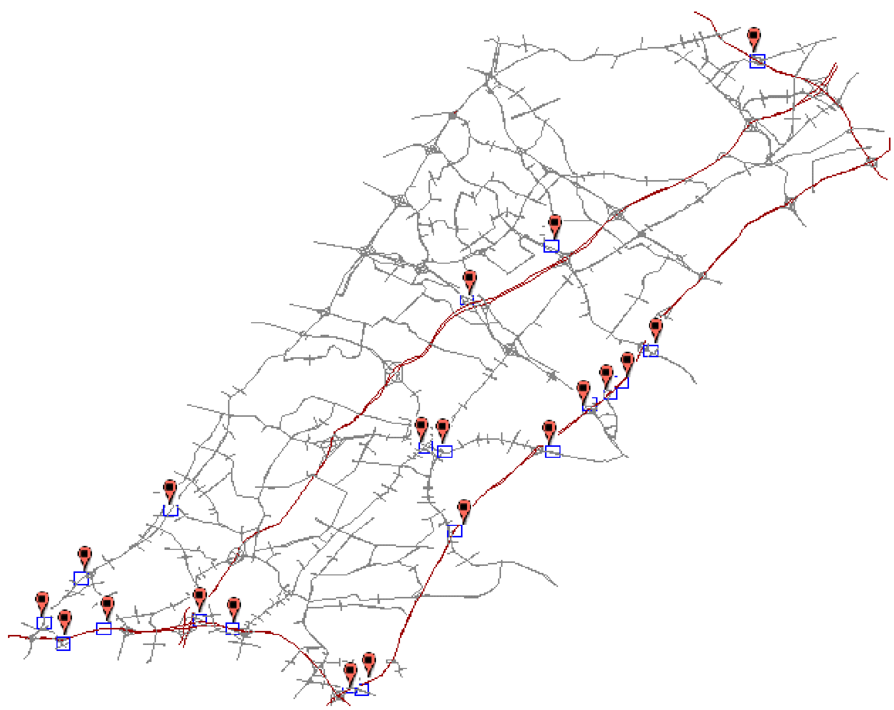


Figure 5-13 Scenario 3 (20 Sensors) Sensor Locations

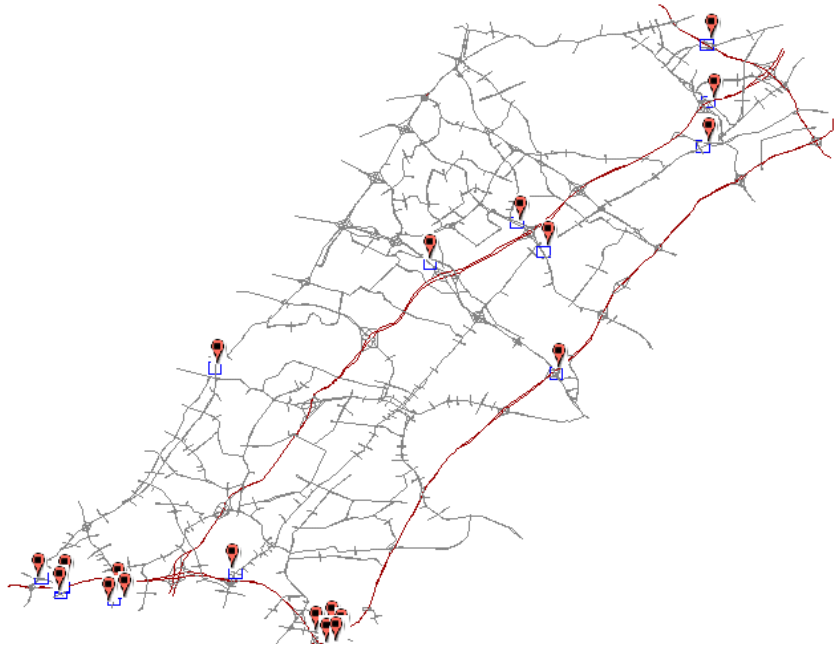


Figure 5-14 Scenario 5 (20 Sensors) Sensor Locations

5.3.2.3.2 Judgment Based Sensor Location Analysis

The second set of sensor location scenarios were generated using engineering judgment to place sensors on high volume links focusing on the main freeways and arterials. This set of scenario analyses should reveal the benefits of adding additional sensor to specific areas in the network. This analysis includes the 9 scenarios described below (each of the scenarios conforms to the filtering criteria):

Scenario 21: top 10 links on I-95 SB and I-95 NB (20 links total)

Scenario 22: top 5 links on I-95 SB, I-95 NB, I-295 SB and I-295 NB (20 links total)

Scenario 23: top 10 links on I-95 SB and I-95 NB and top 5 links on I-295 SB and I-295 NB (30 links total)

Scenario 24: top 5 links on I-95 SB, I-95 NB, I-295 SB, I-295 NB, Rte 1 SB and Rte 1 NB (30 links total)

Scenario 25: top 10 links on I-95 SB and I-95 NB and top 5 links on eastbound and westbound crossroads (30 links total)

Scenario 26: top 5 links on I-95 SB, I-95 NB, I-295 SB, I-295 NB, Rte 1 SB, Rte 1 NB, US29 SB and US29 NB (40 links total)

Scenario 27: top 10 links on I-95 SB and I-95 NB and top 10 links on eastbound and westbound crossroads (40 links total)

Scenario 28: top 10 links on I-95 SB and I-95 NB, top 5 on I-295 SB and I-295 NB, and top 5 links on eastbound and westbound crossroads (40 links total)

Scenario 29: top 10 links on I-95 SB, I-95 NB, I-295 SB, I-295 NB, Rte 1 SB, Rte 1 NB, US29 SB and US29 NB (80 links total)

Figure 5-15 shows a plot of the RMSE for the estimation and prediction for each of the scenarios. The set of scenarios was developed to allow for the exploration of tradeoffs in locating the sensors on different freeways, arterials and crossroads. With this in mind, comparing Scenario 21 (10 links on I-95 SB and I-95 NB) and Scenario 22 (5 links on I-95 SB, I-95 NB, I-295 SB and I-295 NB), both 20 sensor scenarios, one can conclude that locating sensors on I-95 is more valuable than placing them on I-295. This result is consistent with the trends observed in the random selection analysis.

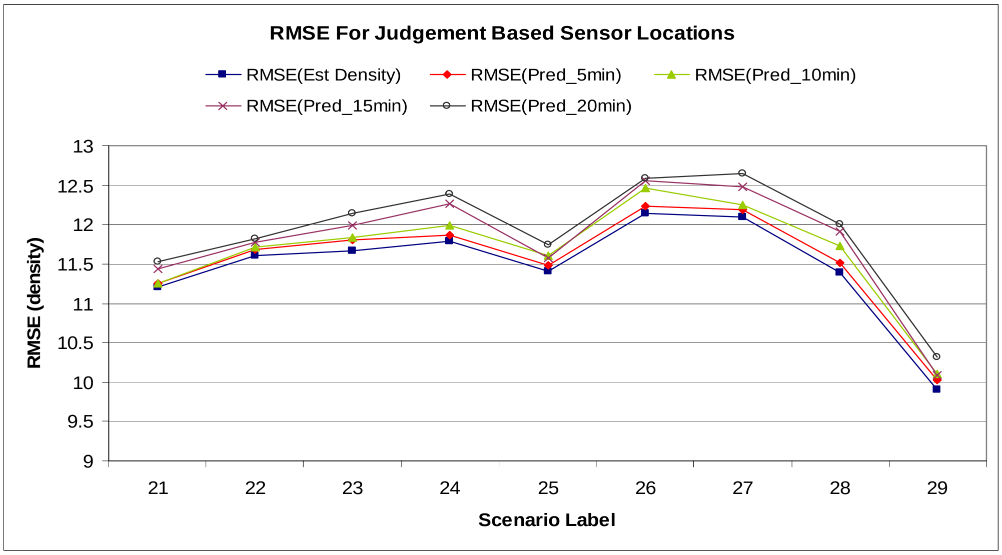


Figure 5-15 RMSE for Judgment Based Sensor Locations

The next sets of comparisons provide less obvious insights. A comparison between Scenario 21 and Scenario 23 or 24 shows that Scenarios 23 and 24 produce no significant changes in performance even though there are more sensors. These additional sensors in this scenario were placed on a much lower volume arterial and sensors could not produce the same level of performance even though there were more of them. A similar result is obtained with Scenario 26, whose performance does not improve over Scenario 25. This result can be attributed to the reduction in sensors on I-95.

Scenarios 22 and 24 are subsets of scenario 26. Scenario 26 performs the worst but has the most detection, while Scenario 22 performs the best and has the least detection. The explanation for this is that the additional sensors have been placed on arterials with much lower volume and the model is in conflict in trying to match both the freeway and arterial sensor information. An approach that can be used to accommodate this conflict and the model's ability to best manage this situation would be to provide a weighting scheme which placed high value on links with higher volume.

Scenario 29 is a scenario with 80 sensors. Scenario 26 is a subset of this scenario. As expected that Scenario 29 outperforms Scenario 26. In this case, critical freeway sensors are added on I-95 and I-295, in addition to the sensors on the minor arterials Route 1 and US 29.

Scenarios 25, 27 and 28 all consider the addition of sensors to east/west crossroads. Again, the results are implying that the addition of sensors on lower volume arterials produces a decline in estimation performance.

Overall, these results suggest that high volume freeways are more valuable as sensor locations than low volume arterials. The analysis also suggests that increasing the number of sensors on freeways is valuable.

5.3.2.3.3 Joint Analysis Results

Looking at the results from both of the analyses, random selection method of sensor location produced lower RMSE. There are a couple of reasons that this may have occurred. First, a random selection of the sensor locations is likely to provide less correlation than the scenarios that were developed based on engineering judgment. The second reason that may have lead to the better performance in the random analysis scenarios is that they included sensors on I-495 and I-695, when the judgment based analysis did not. Freeways I-495 and I-695 are high volume freeways that can greatly influence the estimation performance in the network, and possible the model would have performed better in the judgment scenarios had these freeways not been excluded.

The sensor location problem is a complex optimization problem that can be very difficult to solve, due to the size of the problem and the fact that an optimal solution may not exist. The purpose of the sensitivity analysis of the sensor location and sensor number to the performance of network estimation and prediction is to explore the significance of adding sensors to the network and also to provide insights about the process of selecting the locations for sensors in a network, into the mathematic sensor location model formulation.

5.3.3 SLP Model Experimental Design and Result Analysis

In this section, the proposed mathematic models and their associated HGRASP-DTA heuristic procedures are tested on the CHART network. As explained in last chapter, the deterministic model (DOSLP) is a special case of the stochastic model (SOSLP) under network normal condition. The simulation experiments were implemented on an Intel Xeon CPU 3.20GHZ 64 bits machine with 8G memory. All the algorithms are implemented in Visual Fortran and Visual C++ on the Windows platform with Windows XP professional operation system. The time horizon of interest is the morning peak period from 6:30AM to 8:30AM. As the *a priori* variance and covariance matrix is not available, it is assumed that the *a priori* demand variance is 20% of the demand volume of the corresponding OD pairs in the time-dependent historical demand table. The perturbed time-dependent table is loaded to the simulation software, DYNASMART-P to generate link measurements and time-dependent assignment matrix. The standard deviation of the link flow measurement error is set to 10% of the corresponding simulated link flow.

The HGRASP stopping criterion in this study is set to 10 iterations. The Tabu searching stopping criterion is set to 50 iterations and the Tabu table size is set to 10 links with the Tabu tenure as 2 (the aspiration strategy allows for the revisit of a Tabu move after 2 of the trial moves). The size of the RCL is 364 links that have the highest link flows in the network across the simulation horizon.

In this study, "*stochastic scenario*" is defined as a scenario realization under uncertainty. "*Deterministic scenario*" is defined as a scenario realization under normal (recurrent traffic) conditions. Based on the CHART network incident statistics data in year 2001 and 2002 (table 5-1) (Liu et al. 2004), it is assumed in this study that one or two incidents may occur at the same time during each incident realization. The probability of having one or two incidents in the network would be 0.36 and 0.14 under the assumption of the same link incident occurrence rate $4*10^{-8}$ /veh-lane-mile-day throughout the network. The start time of an incident is 7:00AM and end time is 7:40 AM with severity 0.7, namely the remaining available capacity of the incident link becomes 0.3 or 30 percent of the original link capacity. The impacted traffic diversion rate is assumed to be 80%.

Table 5 - 1 CHART Network Incident Data Collected in Year 2001 and 2002

Available Records		Year 2001		Year 2002	
		Records	Total (%)	Records	Total (%)
CHART II Database	Disabled Veh	16,236	58.6	13,752	41.9
	Incident	8,743	33.6	19,062	58.1
Paper Form (Both Type)		2029	7.8	N/A	N/A
Total		26,008	100	32,814	100

In the proposed HGRASP-DTA solution procedure associated with the SOSLP model, each candidate Tabu move is evaluated under a set of incident realizations and every incident realization requires a single run of the simulation. Apparently, more incident realizations will cause not only better network scenarios representation but also more computational time that is proportional to the network size. A Ranking Similarity Index can be used to compare the solution similarity generated by two different realizations (Chiu et al.. 2001). In this study, we set the incident realization for a candidate sensor location set as 50, which makes the total simulation runs 10*50*50=25,000. In order to balance computational feasibility and solution reliability, it is assumed that impacted vehicles diverted before reaching the incident scene would not affect the vehicles on the alternative routes, given the relatively small portion of impacted vehicles in a large-scale congest network. The vehicle trajectory under normal conditions is considered as the base case; when an incident occurs, the impacted origin and destination zones are delineated. All newly generated vehicles (during the incident) from these impacted origin zones and the en-route impacted vehicles that would have originally traversed the incident link will be classified as user class v^{ω}, provided with diversion guidance to the alternative routes. All other vehicles will be classified as user class o^{ω} and will retain their original assigned paths.

Figure 5-16 shows five most likely incident locations based on the Poisson probability distribution assumption in the Maryland CHART network where three locations (a, c, e) are on I-95 southbound and the other two locations (d, b) on I-495 westbound. Considering the large morning commute traffic volumes from Baltimore to Washington DC and from Maryland to Northern Virginia in the real world, those potential incident locations are reasonable. Interestingly, the existing fourteen detectors in the CHART network depicted in figure 5-7 are deployed mainly along the freeways and in the neighborhood of those most likely incident locations in figure 5-16.

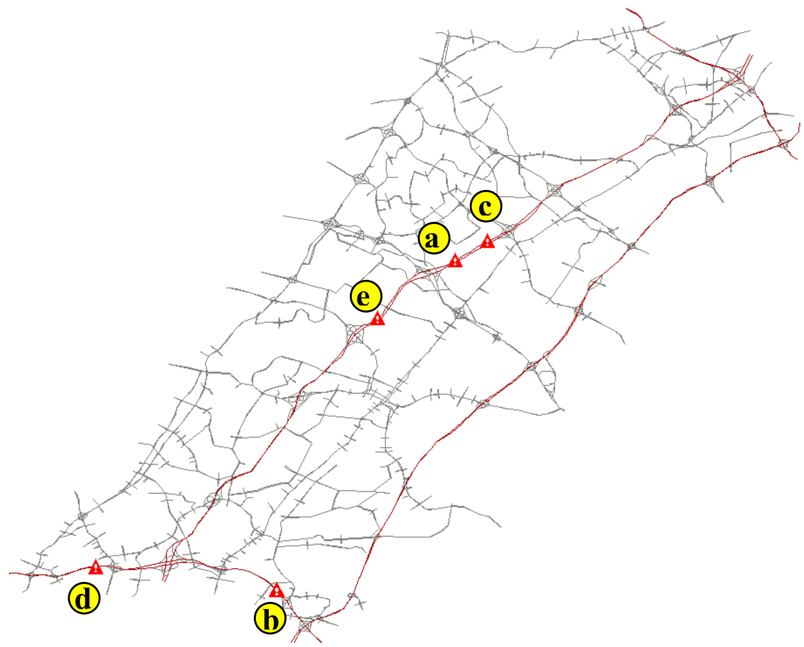


Figure 5-16 Five Most Likely Incident Locations in Maryland CHART Network

Figure 5-17 depicts the zone boundaries and traffic volume among different zones in the CHART network across the two hour (6:30AM-8:30AM) simulation horizon. The width of the blue line that connects origin zone and destination zone is proportional to the OD volumes of the corresponding OD pair.

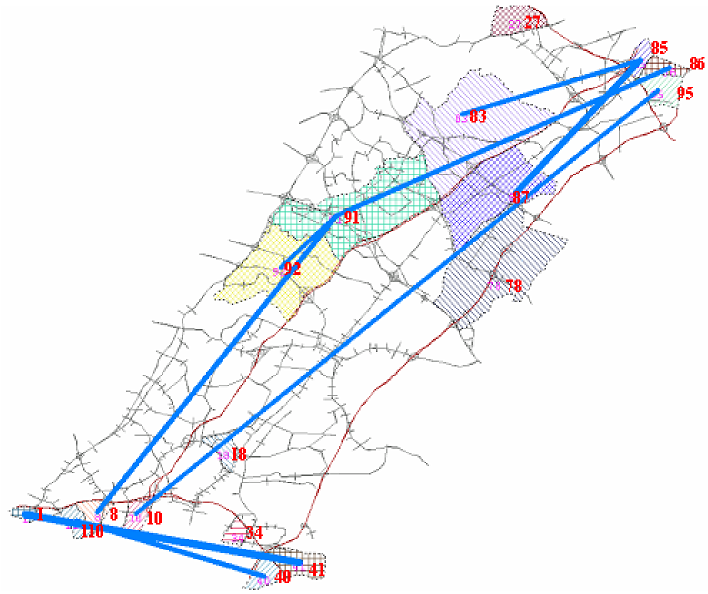


Figure 5-17 Two-Hour Traffic Volume (6:30AM-8:30AM) in CHART Network

5.3.3.1 Effect of the Objective Weight on the Sensor Locations

As discussed earlier, weights on the OD coverage and total link information gains affect the sensor placement, which consequently influence the demand estimation. For this reason, it is essential to understand the effect of the magnitude of the weight on the sensor locations, such that an appropriate weight in the objective function can be determined. The magnitude of the weight λ, the decision maker's preference to the total link information gain, was varied from 0 to 1.0. The effect of varying this weight is shown in table 5-2 where the maximal number of sensors in the network is 30. As demonstrated in table 5-2, the sensitivity of the optimal sensor locations determined by the stochastic model and deterministic model were tested under different scenarios (stochastic scenario and deterministic scenario). The network total OD flow coverage, total link information gains and the associated demand uncertainty reduction were calculated under different scenarios with a variety of weights scaled from 0 to 1. The total uncertainty reduction is calculated using Eq.(5-1)

$$\frac{\sum_t \sum_w P_{w,t}(-) - \sum_t \sum_w P_{w,t}(+)}{\sum_t \sum_w P_{w,t}(-)} \qquad (5\text{-}1)$$

Where $\sum_t \sum_w P_{w,t}(-)$ is the total a *priori* demand variance, $\sum_t \sum_w P_{w,t}(+)$ is the total a *posteriori* demand variance.

Table 5 - 2 OD Coverage and Information Gains for Various Scenarios by 30 Sensors

No Incident (Deterministic Scenario)						
Weights (Link Information Gains, OD Coverage) $(\lambda, 1-\lambda)$	**Stochastic Model Solution**			**Deterministic Model Solution**		
	Total OD Flow Covered	Total Information Gain	Total Demand Uncertainty Reduction	Total OD Flow Covered	Total Information Gain	Total Demand Uncertainty Reduction
(0.0,1.0)	70,756	222.00	12.96%	72,490	232.59	13.90%
(0.2,0.8)	70,186	238.70	14.60%	72,153	272.89	15.97%
(0.4,0.6)	67,624	256.73	15.29 %	72,153	272.89	15.97%
(0.6,0.4)	61,341	276.05	16.43%	71,088	277.17	16.98%
(0.8,0.2)	61,341	276.05	16.43%	71,088	277.17	16.98%
(1.0,0.0)	61,341	276.05	16.43%	69,786	277.47	17.09%
With Incidents (Stochastic Scenario)						
Weights (Link Information Gains, OD Coverage) $(\lambda, 1-\lambda)$	**Stochastic Model Solution**			**Deterministic Model Solution**		
	Expected OD Flow Covered	Expected Information Gain	Expected Demand Uncertainty Reduction	Expected OD Flow Covered	Expected Information Gain	Expected Demand Uncertainty Reduction
(0.0,1.0)	72,430	232.91	13.97%	70,700	221.79	12.92%
(0.2,0.8)	72,074	272.76	15.92%	70,158	239.21	14.68%
(0.4,0.6)	72,074	272.76	15.92%	67,565	256.37	15.21%
(0.6,0.4)	71,004	276.75	16.61%	61,253	275.30	16.26%
(0.8,0.2)	71,004	276.75	16.61%	61,253	275.30	16.26%
(1.0,0.0)	69,705	277.06	16.93%	61,053	276.07	16.49%

As expected, different weight scales result in different location solutions, and link information gain is increased with the augmentation of the weight. Table 5-2 shows that demand coverage improvement and demand uncertainty reduction are two conflicting objectives. Namely, demand coverage percentage is not necessarily proportional to the demand estimation quality. Under normal network conditions, the deterministic model can achieve much more link information gains and flow coverage than the corresponding stochastic model. For example, under normal conditions, total information gain is 277.47 and total flow coverage is 69,786 when $\lambda = 1$ obtained from the deterministic model, while the total information gain is 276.05 and total flow coverage is 61,341 using the stochastic model. Under the stochastic scenario, the expected OD flow coverage and information gains from the stochastic model are greater than those obtained by solving the sensor location problem based on the deterministic model. For example, under stochastic scenario, the expected OD flow coverage is 70,700 and the expected information gain is 221.79 from the deterministic model when $\lambda = 0$, while the expected flow coverage is 72,430 and information gain is 232.91 from the stochastic model. In addition, for the same sensor placement, the deterministic model under deterministic scenario can achieve larger demand uncertainty reduction and OD flow coverage than that under the stochastic scenario. This can be explained by the fact that the deterministic model did not consider the vehicle rerouting during the incident in the formulation, such that it cannot capture the impacted vehicles that took alternative routes when incidents occurred in the network.

From table 5-2 one can also find that the location solutions are not much sensitive to the weights for $\lambda \geq 0.6$ in both models. The likely explanation is that the sensors are more likely located on those links that can intercept OD pairs with large variances when the objective underscores the reduction of demand uncertainty However, the demand uncertainty is assumed to be proportional to the corresponding demand in this study, which gives those links intercepting large OD volumes a higher likelihood of being selected. It also explains why neither of those models is sensitive to the weight when the level of detection is low (i.e, less than 30 sensors in CHART network). A weight $\lambda = 0.6$ is therefore used in the subsequent experiments.

In order to illustrate the weight effect on the sensor placement, figures 5-18 and 5-19 display the optimal sensor location plans obtained from the SOSLP model for 30 sensors when $\lambda = 1$ and $\lambda = 0$. Sensors in figure 5-18 are mainly deployed along the freeways to intercept those OD pairs with large volumes and obtain maximal link information gains. Figure 5-19 shows that sensors (1, 2 and 3) are deployed on the entry/exit links in order to capture the maximal OD flows. Other sensors in figure 5-19 are mostly distributed along the boundary entry/exit links of the OD zones with large demand to provide detection at or around freeway and arterials access points throughout the network. As a commonality, both of the sensor location plans locate the sensors on freeways, arterials and crossroads.

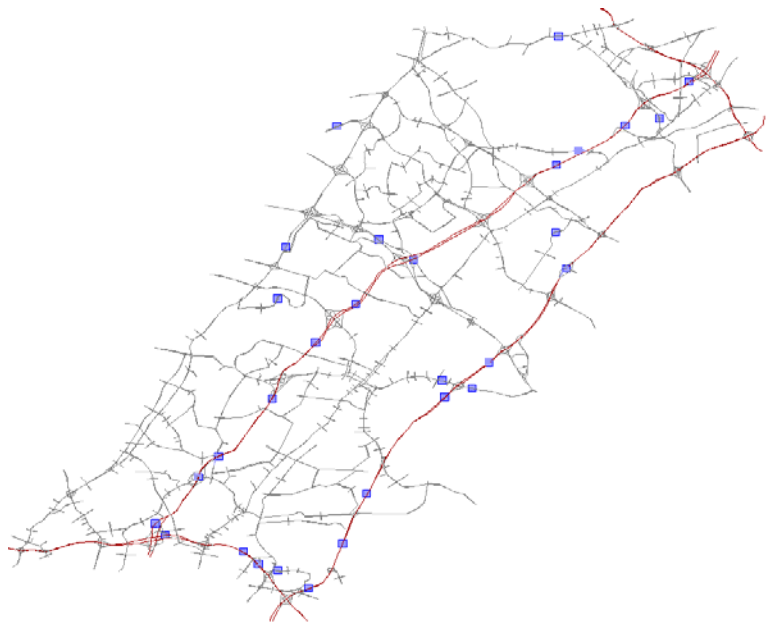


Figure 5-18 30 Sensor Locations by SOSLP model in CHART Network ($\lambda = 1$)

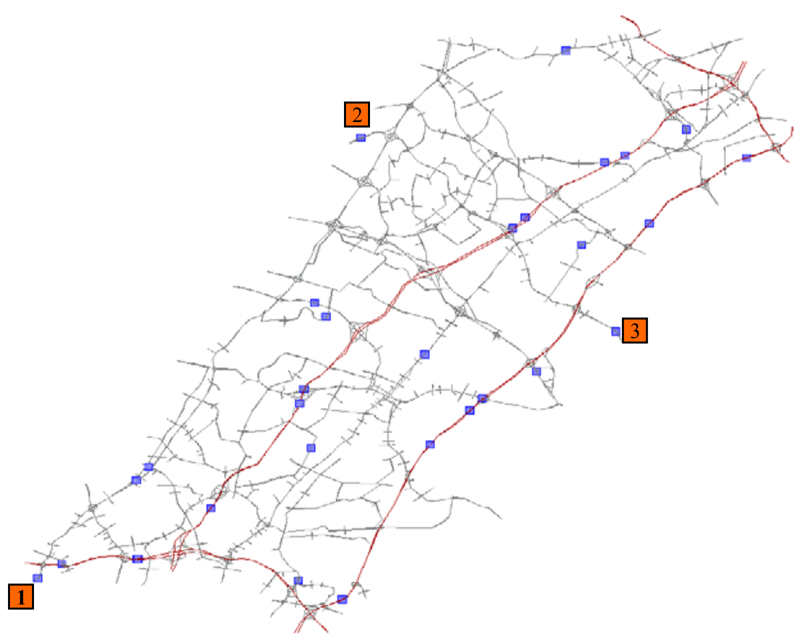


Figure 5-19 30 Sensor Locations by SOSLP model in CHART Network ($\lambda = 0$)

Figure 5-20 illustrates the relationship between the OD pair coverage percentage and weight of 30 sensors obtained from the DOSLP and SOSLP models in stochastic and deterministic scenarios. From the figure, one can observe that 30 sensors can cover at least 50% of the OD flows in the CHART network for both models regardless of the scenarios. In addition, one can also find from the figure that the OD coverage percentages from the stochastic model in stochastic scenario and deterministic model in deterministic scenario are higher than those obtained based on the stochastic model in deterministic scenario and deterministic model in stochastic scenario.

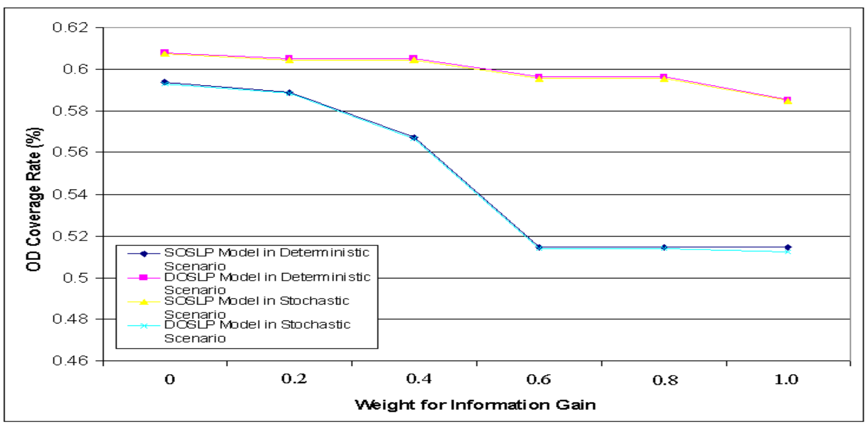


Figure 5-20 O-D Flow Coverage with Different Weight

5.3.3.2 Effect of Sensor Number on the Sensor Locations

This section evaluates the sensor coverage and link information gains of the 10-sensor and 30-sensor plans obtained from the SOSLP model under the stochastic scenario with $\lambda = 0.6$. The principal goal of this section is to demonstrate the marginal value of the newly added sensors in terms of real-time traffic status estimation and prediction.

Table 5-3 denotes the time-dependent demand uncertainty reduction of the 6 highest variances across the OD pairs in the morning peak period from 7:00AM to 8:00AM (12 time intervals) in the stochastic scenario with $\lambda = 0.6$ for the 10- sensor plan. The total number of OD pairs of the CHART network is 12,210 and each OD pair carries less than 1.5% OD flows of the total OD demands of the corresponding time interval. As shown, demands from origin zone 41 to destination zone 1 are the highest across the OD pairs in the study network during most time intervals. In addition, the six OD pairs with the largest demand variance across the OD pairs are covered by at least one sensor in nine time intervals out of 12 total time intervals. Note that although traffic from origin zone 86 to destination zone 91 is not covered during the 7:20AM-7:30AM time interval, it is covered afterwards by the sensor on link 1938, located on I-95 southbound.

Table 5-3 further illustrates that OD pairs associated with large estimation errors are covered by more sensors, which however do not always result in greater uncertainty reduction due to the magnitude of the original uncertainty. For example, OD pair (83, 85) during 7:00AM-7:05AM is covered by three sensors and the uncertainty reduction is 26.59% while the OD pair from origin zone 87 to destination zone 85 during 7:15AM-7:20AM is covered by two sensors and the uncertainty reduction is 37.21%. Table 5-3 also shows that traffic dynamics and time-dependent demand magnitude affect the demand estimation results. For example, the demand uncertainty of OD pair (87, 85) during 7:10AM-7:15AM covered by two sensors was reduced 56.43%, however the uncertainty for the same OD pair covered by the same set of sensors was reduced 37.21% during 7:15 AM-7:20AM.

Table 5 - 3 List of Time-dependent OD Pairs with the 6 Highest Variances of 10 sensors ($\lambda = 0.6$)

Weights ($\lambda = 0.6$)	**Measured Link ID** (56,913,1095,1758,1938,1946,1950,1989,2250,2317)						
Time Interval (Total Demand)	Origin Zone	Dest Zone	Historical 5 –minute Demand	% of Demand To the Total	Posterior Variance	Variance Reduction (%)	# of Sensors Covered
7:00AM-7:05AM (5035 veh/5min)	41	1	58	1.15	120.87	10.18	1
	87	85	55	1.09	60.98	49.61	2
	8	91	55	1.09	79.60	34.21	1
	86	91	45	0.89	59.23	26.88	1
	83	85	45	0.89	59.46	26.59	3
	4	110	45	0.89	76.38	5.70	1
7:05AM-7:10AM (4832 veh/5min)	41	1	72	1.49	193.34	6.76	1
	87	85	54	1.12	37.63	67.73	2
	8	91	51	1.06	78.81	24.25	1
	34	1	41	0.85	64.84	3.57	1
	92	91	40	0.83	64	0.00	0
	83	85	40	0.83	57.38	10.33	3
7:10AM-7:15AM (4739 veh/5min)	87	85	53	1.12	48.95	56.43	2
	86	91	44	0.93	56.01	27.67	1
	41	1	44	0.93	72.97	5.77	1
	78	27	41	0.87	66.63	0.91	2
	8	91	41	0.87	65.86	2.05	1
	40	110	38	0.80	55.16	4.51	1
7:15AM-7:20AM (4684 veh/5min)	41	1	51	1.09	93.70	9.94	1
	8	91	48	1.02	75.34	18.25	1
	88	31	40	0.85	62.37	2.55	1
	86	91	39	0.83	54.71	10.07	1
	40	110	39	0.83	58.47	3.90	1
	87	85	37	0.79	34.39	37.21	2
7:20AM-7:25AM (4858 veh/5min)	41	1	53	1.09	94.84	15.59	1
	87	85	52	1.07	98.69	8.76	2
	8	91	47	0.97	73.55	16.76	1
	86	91	44	0.91	77.44	0.00	0
	95	10	42	0.86	62.55	11.36	2
	40	110	35	0.72	46.16	5.79	1
7:25AM-7:30AM (4914 veh/5min)	41	1	56	1.14	113.67	9.38	1
	87	85	49	1.00	80.49	16.19	2
	40	110	44	0.90	67.46	12.89	1
	83	85	42	0.85	68.16	3.40	3
	86	91	41	0.83	67.24	0.00	0
	34	1	39	0.79	58.91	3.17	1

7:30AM-7:35AM (4955 veh/5min)	41	1	62	1.25	132.29	13.96	1
	87	85	52	1.05	81.83	24.35	2
	40	110	46	0.93	78.28	7.51	1
	8	91	41	0.83	61.44	8.63	1
	83	85	38	0.77	56.10	2.88	3
	86	91	37	0.75	51.85	5.32	1
7:35AM-7:40AM (5036 veh/5min)	41	1	58	1.15	114.88	14.62	1
	86	91	51	1.01	89.00	14.46	1
	8	91	51	1.01	75.47	27.46	1
	87	85	50	0.99	85.65	14.35	2
	83	85	45	0.89	74.93	7.50	3
	34	1	43	0.85	69.11	6.55	1
7:40AM-7:45AM (5079 veh/5min)	41	1	59	1.16	123.66	11.19	1
	8	91	53	1.04	66.08	41.18	1
	87	85	52	1.02	88.58	18.10	2
	95	10	45	0.89	61.28	24.35	2
	86	91	43	0.85	65.34	11.66	1
	83	85	42	0.83	65.26	7.51	3
7:45AM-7:50AM (5037 veh/5min)	41	1	63	1.25	133.58	15.86	1
	8	91	54	1.07	77.41	33.64	1
	87	85	53	1.05	75.48	32.83	2
	86	91	42	0.83	67.73	4.00	1
	40	110	41	0.81	63.78	5.14	1
	34	1	40	0.79	61.24	4.31	1
7:50AM-7:55AM (4973 veh/5min)	87	85	52	1.05	91.38	15.51	2
	41	1	50	1.00	90.66	9.34	1
	86	91	49	0.99	83.24	13.33	1
	8	91	47	0.95	68.32	22.68	1
	34	1	40	0.80	58.62	8.40	1
	43	110	36	0.72	51.44	0.76	1
7:55AM-8:00 AM (4875 veh/5min)	41	1	56	1.15	95.99	23.48	1
	8	91	501	1.03	72.46	27.54	1
	95	10	42	0.86	61.85	12.34	1
	86	91	40	0.82	64.00	0.10	1
	40	110	40	0.82	59.33	7.30	1
	34	1	39	0.80	54.46	10.48	1

* The total number of vehicles in the 2-hour period is 119,189 vehicles

Figure 5-21 shows the optimal 10-sensor location plan obtained based on the SOSLP model for the CHART network with $\lambda = 0.6$. The traffic Analysis Zones (TAZ) with high traffic volumes and the top five most likely incident locations are also displayed in this figure. Sensor 1 on link 913 intercepts the westbound traffic flows on I-495 from origin zone 41 to destination zone 1, which carries the highest volume across the OD pairs in the study network in the morning peak period. Sensor 5 on link 1095 intercepts the eastbound traffic on I-495. Sensors 4,6,7,9 and 10, all located on I-95, intercept the northbound and southbound traffic. Sensor 2 is located on US-29. Note that sensors are mainly located along I-95, and no sensors are on I-295 in figure 5-21. This result reaffirms the earlier finding in section 5.3.2 that locating sensors on I-95 is more valuable than placing them on I-295 when the budget is constrained.

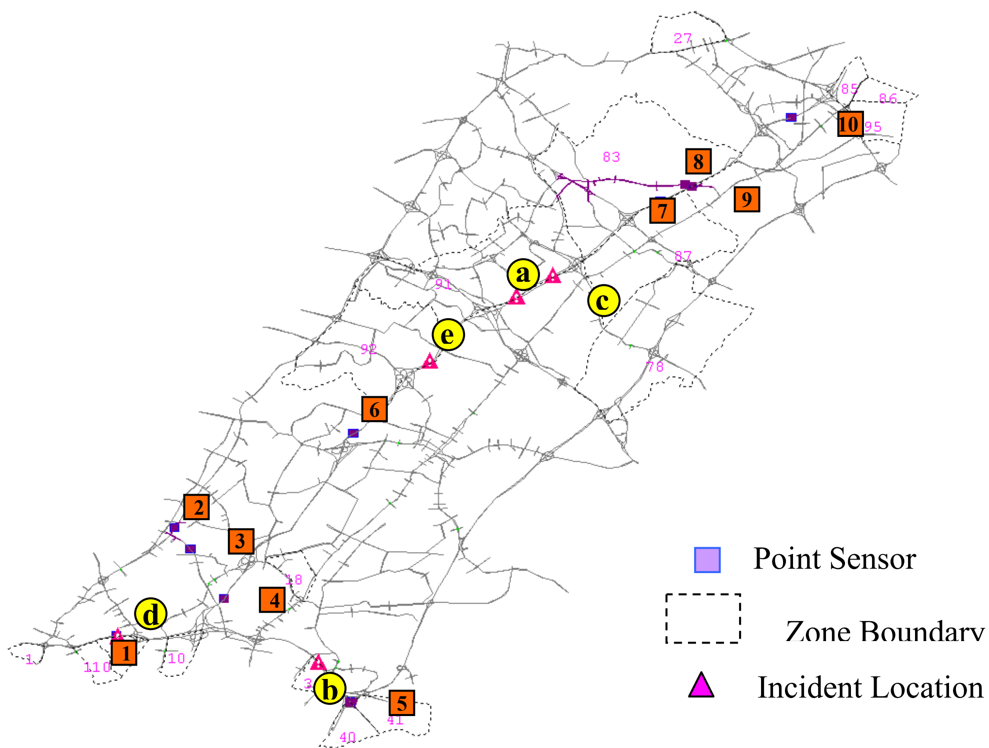


Figure 5-21 10 Sensor Location Plan Obtained from SOSLP in CHART Network

Table 5-4 displays the time-dependent demand uncertainty reduction with the six highest variances across the OD pairs in the morning peak period from 7:00AM to 8:00AM under the stochastic scenario with $\lambda = 0.6$ for the 30-sensor location plan obtained from the SOSLP model. Compared to the results of the 10-sensor plan in table 5-3, additional new sensors in this plan covered more OD pairs. Meanwhile, OD pairs associated with a large variance in the network are covered by additional new sensors, which result in a significant improvement in the demand estimation quality. For example, all of the six OD pairs with the largest demand variance across the OD pairs are covered by at least one sensor in 10 out of 12 time intervals. While OD pair (92, 91) was not covered by a sensor in table 5-3 during time 7:05AM-7:10AM, it is covered by two sensors in table 5-4, resulting in 38.74% uncertainty reduction. For OD pair (41,1), which carries the largest OD volumes across all the OD pairs in most of the time intervals, it was covered by one sensor in table 5-3 and the uncertainty reduction was less than 24%. However, it is covered by 3 sensors in table 5-4 and a significant improvement (over 50% uncertainty reduction) is obtained during each time interval, with the largest reduction (89.46%) occurring during the 7:35AM-7:40AM interval. In addition, table 5-4 shows that OD pair (95, 10) is covered by 7 and 8 sensors respectively during 7:20AM-7:25AM and 7:40AM-7:44AM, resulting in 100% uncertainty reduction. Note that although the 30-sensor plan leads to a significant improvement in the demand uncertainty reduction compared to the 10-sensor plan, OD pair (86, 91) is still not covered by any sensor during time interval 7:20AM-7:30AM in table 5-4.

Table 5 - 4 List of Time-dependent OD Pairs with the 6 Highest Variances of 30 sensors ($\lambda = 0.6$)

Weights ($\lambda = 0.6$)	**Measured Link ID** (37,48,308,426,526,764,952,967,1051,1258,1267,1373,1446,1732,1853,1859, 1863,1874,1887,1898,1950,1989,2036,2126,2199,2237,2252,2504,2522,2547)						
Time Interval (Total Demand)	Origin Zone	Dest Zone	Historical 5 –minute Demand	% of Demand To the Total	Posterior Variance	Variance Reduction (%)	# of Sensors Covered
7:00AM-7:05AM (5035 veh/5min)	41	1	58	1.15	72.22	46.33	3
	87	85	55	1.09	88.81	26.60	1
	8	91	55	1.09	34.87	71.18	2
	86	91	45	0.89	49.77	38.55	2
	83	85	45	0.89	60.60	25.18	2
	4	110	45	0.89	65.16	19.55	3
7:05AM-7:10AM (4832 veh/5min)	41	1	72	1.49	32.81	84.18	3
	87	85	54	1.12	75.71	35.09	1
	8	91	51	1.06	63.17	39.28	2
	34	1	41	0.85	46.72	30.52	3
	92	91	40	0.83	39.20	38.74	2
	83	85	40	0.83	58.58	8.46	2
7:10AM-7:15AM (4739 veh/5min)	87	85	53	1.12	79.06	29.64	1
	86	91	44	0.93	43.93	43.27	2
	41	1	44	0.93	33.49	56.75	3
	78	27	41	0.87	66.93	0.46	1
	8	91	41	0.87	37.03	44.93	2
	40	110	38	0.80	34.71	39.92	3
7:15AM-7:20AM (4684 veh/5min)	41	1	51	1.09	46.70	55.11	3
	8	91	48	1.02	19.17	79.20	2
	88	31	40	0.85	34.19	46.58	5
	86	91	39	0.83	50.71	16.65	2
	40	110	39	0.83	41.60	31.63	3
	87	85	37	0.79	43.79	20.04	1
7:20AM-7:25AM (4858 veh/5min)	41	1	53	1.09	45.04	59.91	3
	87	85	52	1.07	105.81	2.17	1
	8	91	47	0.97	65.85	25.47	2
	86	91	44	0.91	77.44	0.00	0
	95	10	42	0.86	0.01	99.89	7
	40	110	35	0.72	36.75	25.01	3
7:25AM-7:30AM (4914 veh/5min)	41	1	56	1.14	52.04	58.52	3
	87	85	49	1.00	90.42	5.85	1
	40	110	44	0.90	57.44	25.82	3
	83	85	42	0.85	69.43	1.61	2
	86	91	41	0.83	67.24	0.00	0
	34	1	39	0.79	42.15	30.72	3

7:30AM-7:35AM (4955 veh/5min)	41	1	62	1.25	74.53	51.53	3
	87	85	52	1.05	96.12	11.13	1
	40	110	46	0.93	57.38	32.21	3
	8	91	41	0.83	27.08	59.73	3
	83	85	38	0.77	56.62	1.97	2
	86	91	37	0.75	50.51	7.77	2
7:35AM-7:40AM (5036 veh/5min)	41	1	58	1.15	14.19	89.46	3
	86	91	51	1.01	80.61	22.52	2
	8	91	51	1.01	61.67	40.72	2
	87	85	50	0.99	93.38	6.62	1
	83	85	45	0.89	78.16	3.52	2
	34	1	43	0.85	51.43	30.47	3
7:40AM-7:45AM (5079 veh/5min)	41	1	59	1.16	39.25	71.81	3
	8	91	53	1.04	56.21	49.97	1
	87	85	52	1.02	98.28	9.13	1
	95	10	45	0.89	0.00	100.00	8
	86	91	43	0.85	61.23	17.21	2
	83	85	42	0.83	67.46	4.39	2
7:45AM-7:50AM (5037 veh/5min)	41	1	63	1.25	51.61	67.49	3
	8	91	54	1.07	81.79	29.87	3
	87	85	53	1.05	95.99	14.56	1
	86	91	42	0.83	67.73	4.01	2
	40	110	41	0.81	50.88	24.33	3
	34	1	40	0.79	46.20	27.81	3
7:50AM-7:55AM (4973 veh/5min)	87	85	52	1.05	103.29	4.50	1
	41	1	50	1.00	47.77	52.22	3
	86	91	49	0.99	79.64	17.07	2
	8	91	47	0.95	23.01	73.95	2
	34	1	40	0.80	43.66	31.78	3
	43	110	36	0.72	36.59	29.42	4
7:55AM-8:00 AM (4875 veh/5min)	41	1	56	1.15	56.66	54.83	3
	8	91	501	1.03	34.74	65.26	2
	95	10	42	0.86	15.05	78.67	7
	86	91	40	0.82	62.57	2.24	2
	40	110	40	0.82	46.65	27.11	3
	34	1	39	0.80	40.84	32.88	3

* The total number of vehicles in the 2-hour period is 119,189 vehicles

The results in tables 5-3 and 5-4 indicate that significant improvements in uncertainty reduction could be attained by deploying additional new sensors into the network to intercept more OD flows. Unfortunately, although some OD pairs associated with large variances in table 5-3, such as (8, 91) and (40,110) have been covered by additional new sensors in table 5-4, they still have significant associated uncertainty due to the magnitude of the original uncertainties. It is imperative to help transportation planners decide whether to deploy new sensors so as to cover those unobserved OD pairs that may have small variances or continue focusing on those covered OD pairs that still have large variances. The next section aims to characterize the marginal value of the newly added sensors in terms of demand estimation and flow coverage.

5.3.3.3 Sensor Marginal Value

The sensor location problem is viewed in this study from the perspective of the value of information. Sensors continuously provide information that helps characterize the status of the network The key question here is how to characterize the marginal value from a newly added sensor in the context of traffic status estimation and prediction. A sensitivity analysis is conducted to explore the relationship between the number of sensors and level of OD coverage as well as between the number of sensors and level of demand uncertainty reduction in the network. The purpose of this analysis is to explore the marginal value, in terms of flow percentage coverage and demand uncertainty reduction, of adding sensors to the network. The analysis also provides a platform to investigate the effect of sensor location on the OD demand coverage rate.

Tables 5-5 and 5-6 list the expected/total OD coverage, expected/total link information gain and expected/total uncertainty reductions for different number of sensors in the network with $\lambda = 0.6$ obtained from DOSLP and SOSLP models in stochastic and deterministic scenarios. The uncertainty reductions are calculated using Eq.5-1.

Table 5 - 5 Statistics of Different Optimal Sensor Location Plans in Stochastic Scenario ($\lambda = 0.6$)

Problem Size $G\ (V, A)$	G(2182,3387)			
	With Incidents (Stochastic Scenario)			
Sensor Plan (Stochastic Model Solution)	Expected OD Flow Covered	(%) Network OD Coverage	Expected Information Gain	(%) Expected Uncertainty Reduction
5 Sensors	31,295	26.26%	68.27	6.69 %
10 Sensors	47,132	39.54%	96.45	7.42 %
15 Sensors	47,809	40.11%	156.60	11.07 %
20 Sensors	62,988	52.85%	180.43	12.85%
25 Sensors	59,776	50.15%	244.86	15.91%
30 Sensors	71,004	59.97%	276.75	16.61%
35 Sensors	66,662	55.93%	366.34	25.33 %
40 Sensors	68,476	57.45%	407.54	27.18 %
45 Sensors	75,202	63.09%	408.13	27.36%
	With Incidents (Stochastic Scenario)			
Sensor Plan (Deterministic Model Solution)	Expected OD Flow Covered	(%) Network OD Coverage	Expected Information Gain	(%) Expected Uncertainty Reduction
5 Sensors	28,616	24.01%	52.13	2.49%
10 Sensors	44,662	37.47%	93.11	6.28%
15 Sensors	52,618	44.15%	127.81	9.91%
20 Sensors	57,963	48.63%	174.95	12.05%
25 Sensors	59,302	49.75%	238.38	15.36%
30 Sensors	61,253	51.39%	275.30	16.26%
35 Sensors	61,762	51.82%	318.99	18.20 %
40 Sensors	65,506	54.96%	398.90	26.45%
45 Sensors	72,522	60.85%	403.06	27.15%

* The total number of vehicles in the 2-hour period is 119,189 vehicles

Table 5 - 6 Statistics of Different Optimal Sensor Location Plans in Deterministic Scenario ($\lambda = 0.6$)

Problem Size $G\ (V, A)$	G(2182,3387)			
	No Incident (Deterministic Scenario)			
Sensor Plan (Stochastic Model Solution)	Total OD Flow Covered	(%) Network OD Coverage	Total Information Gain	(%) Total Uncertainty Reduction
5 Sensors	31,406	26.35%	68.51	6.71%
10 Sensors	47,233	39.63%	96.79	7.43%
15 Sensors	47,885	40.18%	156.81	11.08%
20 Sensors	63,042	52.89%	180.28	12.84%
25 Sensors	59,860	50.22%	235.27	14.91%
30 Sensors	61,341	51.47%	276.05	16.43%
35 Sensors	66,715	55.97%	316.64	17.33%
40 Sensors	68,509	57.48%	398.33	26.40%
45 Sensors	72,234	60.60%	406.18	27.34%
	No Incident (Deterministic Scenario)			
Sensor Plan (Deterministic Model Solution)	Total OD Flow Covered	(%) Network OD Coverage	Total Information Gain	(%) Total Uncertainty Reduction
5 Sensors	28,674	24.06%	72.09	7.49%
10 Sensors	44,780	37.57%	103.09	8.28%
15 Sensors	52,710	44.22%	128.11	9.93%
20 Sensors	57,985	48.65%	184.93	13.04%
25 Sensors	60,373	50.65%	239.02	15.38%
30 Sensors	71,088	59.64%	277.17	16.98%
35 Sensors	71,790	60.23%	318.29	18.15%
40 Sensors	75,519	63.36%	399.71	26.45%
45 Sensors	75,555	63.39%	429.11	28.09%

* The total number of vehicles in the 2-hour period is 119,189 vehicles

As expected, more sensors cover more O-D flows in both scenarios. The inclusion of additional link flow observations determined through optimally- selected additional sensor locations improves the precision of the estimated trip matrix in that the *posteriori* O-D demand variance is reduced.

Table 5-5 shows that, in the stochastic scenario, the SOSLP model achieved larger improvement in demand uncertainty reduction and flow coverage than those obtained based on the DOSLP model. Similarly, table 5-6 demonstrates that, in the deterministic scenario, the deterministic model performs better than the stochastic model in terms of demand coverage and uncertainty reduction. The results from table 5-5 and 5-6 conffirm the effect of the traffic dynamics on the sensor locations. The marginal value of information from additional sensors can be characterized in terms of the demand coverage increase rate and uncertainty reduction rate. In addition, the results show that, in the stochastic scenario, the marginal reduction in uncertainty due to an additional sensor in the stochastic model is not significantly greater than that obtained based on the deterministic model.

Note that in table 5-5, the OD coverage by 30 sensors under the stochastic scenario is 71,004, which is greater than that covered by 35 sensors (66,662). However, the information gain of the 30 sensor plan is 276.75, which is smaller than that from the 35-sensor plan (366.34). It illustrates the fact that maximization of the sensor network coverage does not necessarily result in the largest improvement of the overall OD demand estimation quality.

Figure 5-22 plots the information gains obtained from the different sensor location plans. It shows that the link information gain obtained from the stochastic model in the stochastic scenario is the largest in all four cases while the link information gain obtained from the deterministic model in stochastic scenario is the smallest. Figure 5-23 illustrates the relationship between the number of sensors and the O-D flow coverage rate. It confirms the previous finding in section 5.2 that obtaining greater than 50% OD coverage of the network requires a significant increase in the number of sensors. It shows that 20 sensors covered around 50% of the O-D flows and 45 sensors covered 60% O-D flows. Both figures 5-22 and 5-23 show that the sensors' marginal value is reduced in terms of the flow coverage rate or demand uncertainty reduction when more sensors are deployed into the network.

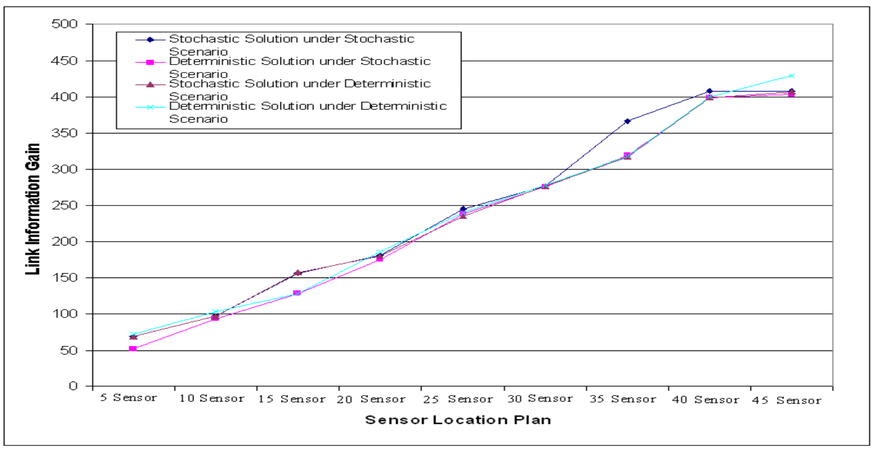


Figure 5-22 Information Gain for Different Sensor Location Plan

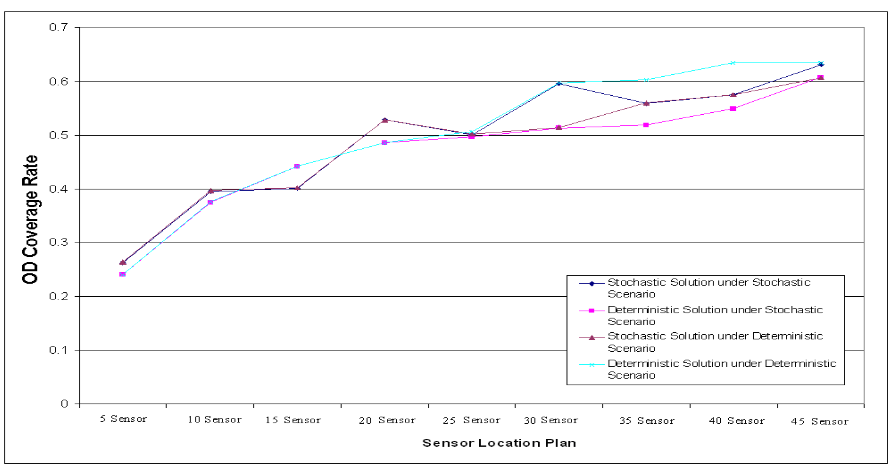


Figure 5-23 OD Flow Coverage for Different Sensor Location Plan

Figures 5-24 and 5-25 show the different optimal sensor location plans obtained from the SOSLP and DOSLP models. The figures indicate that ,more sensors were deployed along the I-95 and I-495 according to the SOSLP model in the stochastic scenario than under the DOSLP model in the deterministic scenario. This is explained by the high incident probability along these two freeways due to the large OD volume. For example, under sensor location plan (a), in figure 5-24, three sensors are deployed along I-95 and two sensors along I-495, while in the corresponding plan in figure 5-25, there are two sensors on I-95 and two sensors on I-295. In addition, more sensors in figure 5-24 are deployed on or close to the freeways and main arterials compared to the those obtained based on the DOSLP model in figure 5-25.

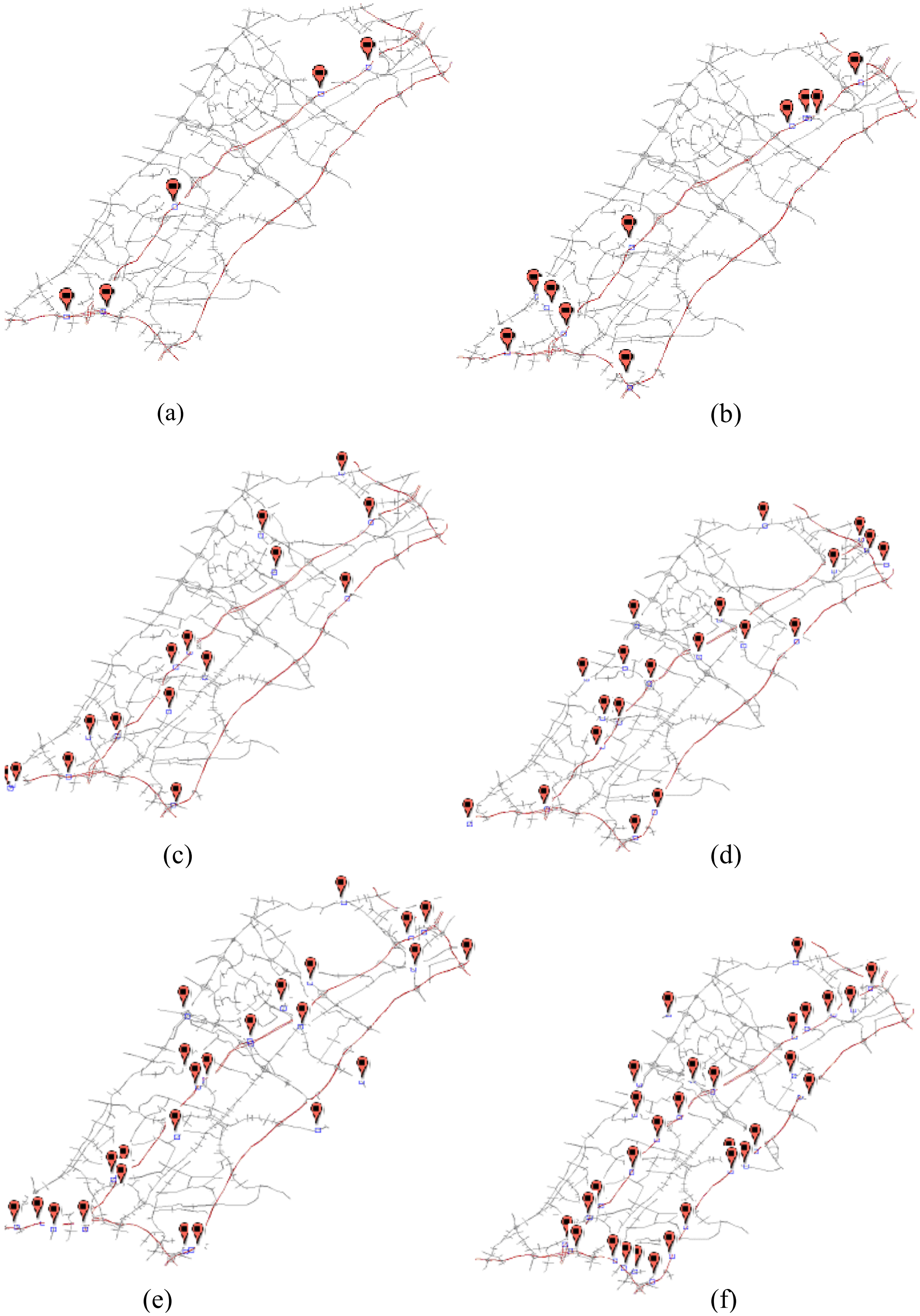


Figure 5-24 Sensor Locations Plan for (a) 5 Sensors, (b) 10 Sensors, (c) 15 Sensors, (d) 20 Sensors, (e) 25 Sensors, (f) 30 Sensors from SOSLP Model

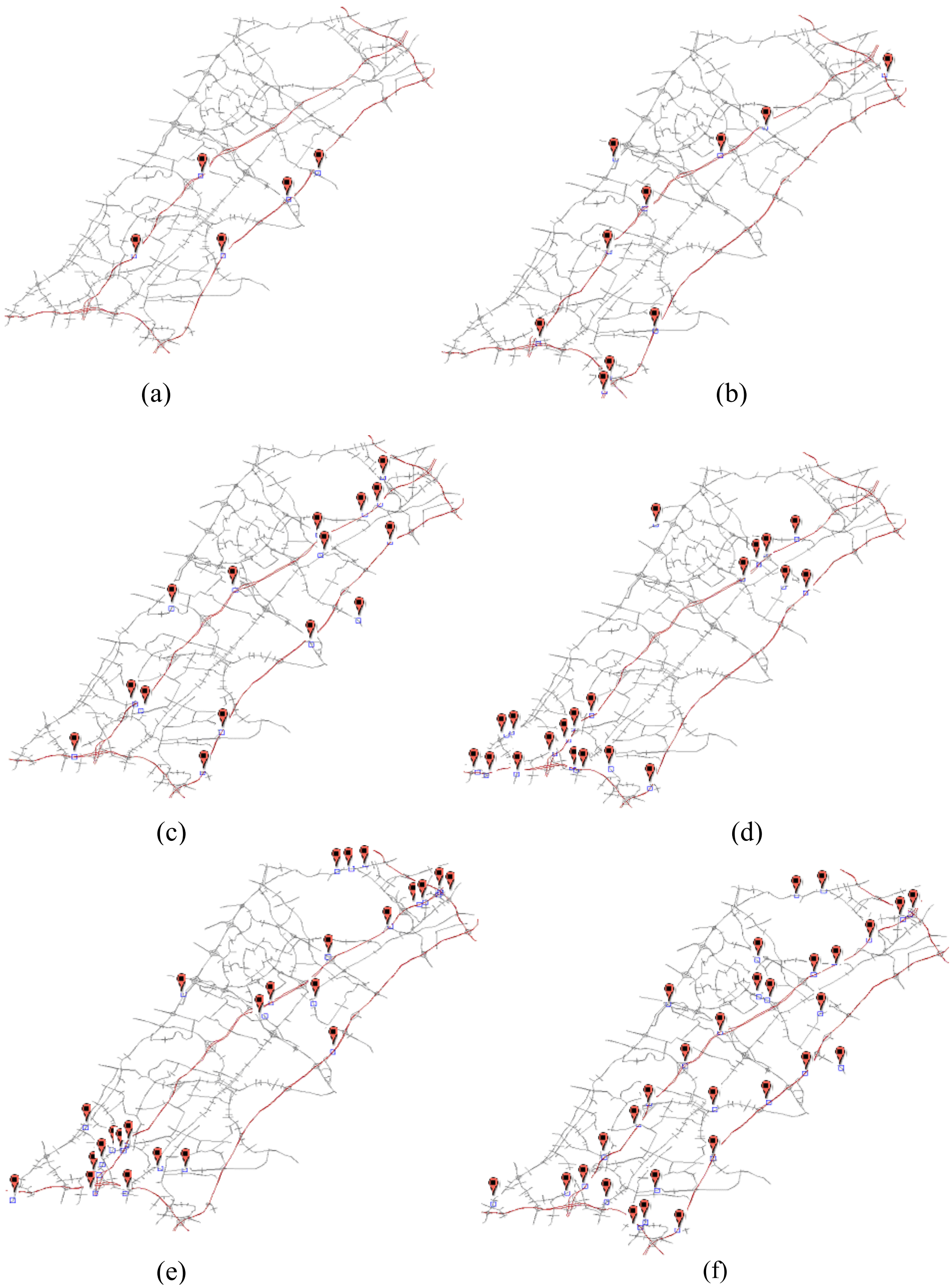


Figure 5-25 Sensor Locations Plan for (a) 5 Sensors, (b) 10 Sensors, (c) 15 Sensors, (d) 20 Sensors,(e) 25 Sensors, (f) 30 Sensors from DOSLP Model

5.3.4 Robustness Analysis with Real-Time OD Estimation and Prediction

This section implements the robustness analysis of different sensor location plans obtained from the SOSLP model to evaluate the performance of the proposed stochastic model under different degrees of real-time information availability, and characterize the marginal value of newly added sensors in terms of traffic status estimation and prediction.

For the experiments conducted in this section, limited real-time field data were available. Therefore the experimental data that is used to mimic real-time sensor information was synthesized using a dynamic traffic simulator, DYNASMART-P. A historical time-dependent matrix corrected by the actual link counts is treated as the "ground truth" for experimental purposes. The ground truth OD demand is loaded onto the network using DYNASMART-P to generate both link counts and density (simulated link measurements). The values become the "sensor data" or "observations" in the synthetic data set. The sensor data served as input to a real-time dynamic traffic assignment package, namely DYNASMART-X (Mahmassani et al. 1998) to evaluate the different sensor plans performance in the network estimation and prediction.

Another input to the procedure is the *a priori* OD demand matrix. It was obtained in this case by perturbing the "ground truth" matrix—assuming it was 80% under-estimated. This *a priori* OD matrix is then combined with the sensor data (from the ground truth simulation) for real-time traffic estimation and prediction. Table 5-7 summarizes the scheduling parameters of DYNASMART-X applied in the experiment. It defines the module execution frequency and length as well as the observation sampling frequency.

Table 5 - 7 System Scheduling Parameters

	Parameter	Value
General parameters	Assignment Interval	5 min
	Observation Interval for LTCC or STCC	5 min
	Observation Interval for ODEC	5 min
Module parameters	RT-DYNA Roll Period	0.5 min
	P-DYNA Roll Period	5 min
	P-DYNA Prediction Horizon	20 min
	ODE State Length	5 min
	ODP Execution Cycle	10 min
	ODP Prediction Horizon	45 min
	Long Term Consistency Checking Cycle (LTCC Period)	5 min
	Short Term Consistency Checking Cycle (STCC Period)	5 min

The *a priori* link estimation density is generated by loading the *a priori* (perturbed OD matrix in this study) demand onto the network, while the online link estimation density is the real-time dynamic traffic assignment results by integrating the (true) real-time link observation data with the *a priori* demand into the estimation. In order to interpret the influence that different sensor plans have on the ability to estimate link-level traffic states, the Root Mean Squared Error (RMSE) in link density is selected as the performance measure:

$$RMSE = \sqrt{\frac{\sum_{l,t}\left(C_{l,t} - C'_{l,t}\right)^2}{n_{obs}}} \tag{5-2}$$

where,

$C_{l,t}$ = observed density on link l during time interval t (ground truth output)

$C'_{,l,t}$ = simulated density on link l during time interval t (simulated output)

n_{obs} = Number of total observations

Table 5-8 depicts the network average link density RMSE with 1 minute observation time interval. The existing 14-sensor location plan serves as the benchmark to compare the effects of optimal sensor location plans on the traffic estimation. It shows that as more sensors are deployed into the network, the estimation error is monotonically reduced. These results demonstrate that optimally deployed sensors could improve the network state estimation quality when utilizing the on-line estimator.

Table 5 - 8 Network Average Link Density RMSE with Different Optimal Sensor Location Plan

Density / Sensor Plan	A *priori* Link Density Estimation	Online Link Density Estimation	(%)Percentage Improvement
5 Sensors	28.50	21.37	25.02%
10 Sensors	28.50	20.84	26.88%
Exiting 14 Sensors	28.50	20.73	27.26%
20 Sensors	28.50	20.63	27.61%
30 Sensors	28.50	20.50	28.07%
40 Sensors	28.50	19.91	30.14%

Figure 5-26 shows the estimated link density on link 1778 (figure 5-7) to further illustrate the effect of different sensor locations on the network state estimation quality. As expected, the online density estimation exhibits a slower changing pattern than the corresponding observation value. In addition, comparing to the a *priori* estimation, the 20-sensor location plan can recognize and capture the density changes on the link.

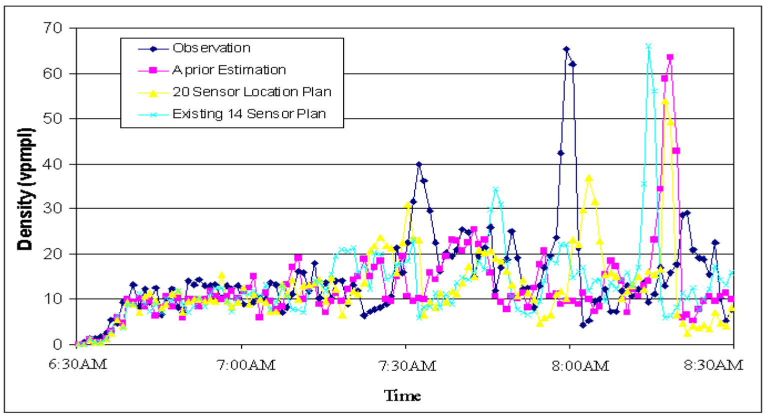


Figure 5-26 Estimated Link Density on Link 1778

The time-dependent average network link density RMSE at 5 minute time intervals are plotted in figure 5-27. It shows that the sensors can reduce average link density errors. Moreover, the additional information from newly added sensors can improve the quality of network traffic status estimation. In addition, Figure 5-27 illustrates that the estimation errors of the link density are proportional to the network congestion level.

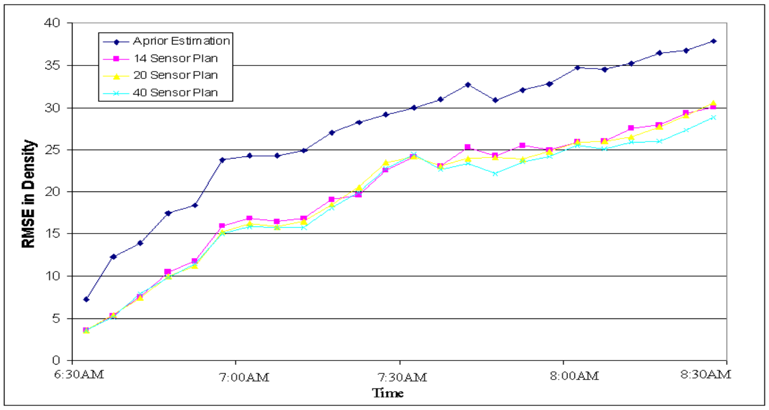


Figure 5-27 Time-Dependent Average Link Density RMSE

In order to interpret the influence of different sensor plans to the accuracy of the estimated OD demand, the RMSE at 5 minute time interval in terms of the time-dependent OD demand (6:30AM-8:30AM) is selected as the performance measure:

$$RMSE_t = \sqrt{\frac{\sum_w \left(d_{w,t} - d'_{w,t}\right)^2}{n_{od}}} \qquad (5\text{-}3)$$

where,

$d_{w,t}$ = Ground-Truth demand of OD pair w during time interval t

$d'_{w,t}$ = Estimated demand of OD pair w during time interval t (simulated output)

n_{od} = Number of OD pairs

Figure 5-28 plots the time-dependent RMSE of different sensor location plans. One can observe from the figure that the estimation errors of the demand and the fluctuation of the error decrease with the newly added sensors in the network.

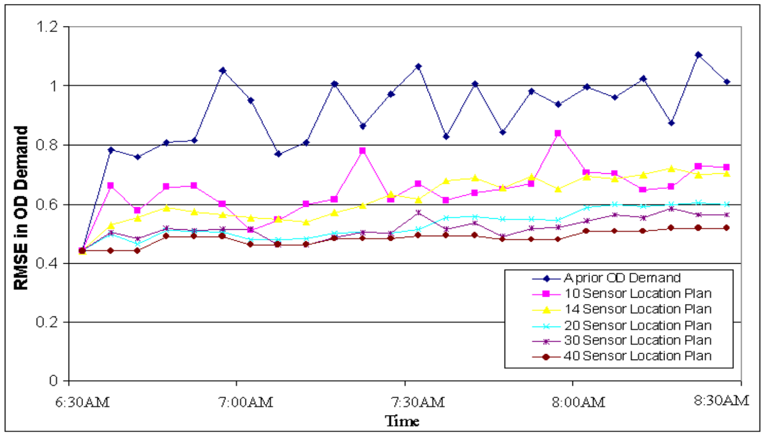


Figure 5- 28 Time-Dependent Demand RMSE in Different Sensor Location Plan

The average demand RMSE is plotted in Figure 5-29 for different sensor plans. In a given sensor location plan, the RMSE is calculated across all the OD pairs and across all of the time intervals. As expected, more optimally deployed sensors lead to greater demand estimation error reduction.

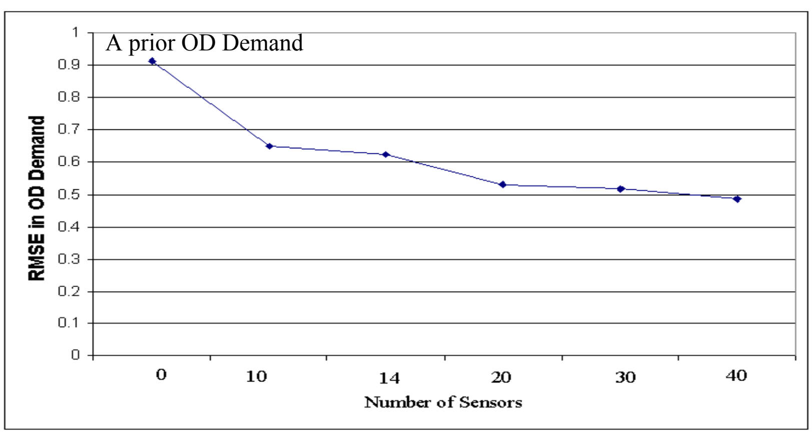


Figure 5- 29 Demand RMSE in Different Sensor Location Plan

5.4 Summary

Increasingly, sensors or detectors are being deployed to monitor network conditions. Installing and maintaining sensors in a transportation network can be expensive. This chapter explores ways to efficiently allocate resources so as to generate a network detection system in a manner that produces minimal estimation errors and minimal equipment costs.

The sensor location problem is interpreted as a value of information problem, which leads to interpretation with learning process models. The analysis provided several valuable insights about the process of selecting the locations for sensors in a network. The difficulty of determining best locations on the basis of judgment alone is an important caveat to learn from the study. A second valuable result was the particular emphasis on the improvement added by placing sensors on high volume freeways. Numerical experiments based on two medium size networks, Fort-Worth, TX and Irvine, CA are used to demonstrate the relationship between the OD coverage and sensor number based on dynamic traffic assignment methodology.

The SOSLP model is an extension of the DOSLP model by considering the network uncertainty in conjunction with the essential impacted vehicles in the non-recurrent congestion. A large-scale network, Baltimore-Washington network (CHART) is used to illustrate the proposed DOSLP and SOSLP models and their associated HGRASP-DTA solution procedures under normal and uncertain traffic conditions. The effect of the magnitude of the weight on the link information gain is investigated. The network performance of different optimal sensor location plans obtained based on DOSLP and SOSLP model is analyzed. Sensitivity analysis on the weight and sensor location is used to evaluate the robustness of the proposed methodologies. In addition, a series of experiments helps to characterize the marginal value of newly added sensor in traffic status estimation and prediction. The estimated demand RMSE analysis demonstrates that sensor's number and location play a critical role in the demand estimation quality.

In summary, the experiments provide and confirm the following important findings for the sensor location problem: (1). The sensors need to be located on the links that can intercept the most OD flows; (2). The sensor observation data should be linearly independent; (3). More sensors do not necessarily mean larger information gains; (4). The lower the measurement error, the more gains the system can obtain; (5). Maximization of the sensor network coverage does not necessarily make the largest improvement in the overall OD demand estimation quality.

The next chapter provides an overall summary and conclusions to this research. The research contributions and future possible extensions of the sensor location problem are highlighted.

Chapter 6 Conclusions and Future Research

This chapter summarizes the research work presented in this dissertation. Section 6.1 draws overall conclusions regarding the proposed framework, methodology and findings. Section 6.2 presents the author's perspective on the contribution to the sensor location problem for real-time traffic estimation and prediction in large-scale networks. Section 6.3 discusses further extensions and outlines several directions for future research in this area.

6.1 Overall Conclusions

Origin-destination (OD) demand is an important input to various transportation network modeling problems. Substantial research has been conducted on demand estimation and prediction to obtain reliable demand for urban traffic networks. However, all existing demand estimation approaches were implemented under the assumption of given sensor locations. Hence the motivation for research on reducing the demand uncertainty through optimal sensor deployment in a network.

The sensor location problem is a complex optimization problem that could be difficult to solve because of its size, and the fact that an optimal solution may not exist. The existing limited body of research on the sensor location problem is mainly focused on flow capture and OD coverage under the assumption of a static traffic flow pattern. With a heavy emphasis on the OD demand estimation problem, this dissertation presents a framework that takes demand estimation and sensor location simultaneously into account based on dynamic traffic assignment (DTA) methodology. It views the sensor location problem as a traffic status learning process that needs sensors to add valuable information that can be used to update estimates (in terms of mean and variance) of the network traffic status in conjunction with the time-dependent demand coverage.

This dissertation starts from a deterministic optimal sensor location problem (DOSLP). It discusses demand coverage of different sensor plans obtained from the DOSLP model and their subsequent effect on the estimation quality of the network state under two situations, with and without budgetary constraints. Because uncertain events, such as incidents, natural disasters, etc, may impact the vehicle paths and the associated traffic pattern, the DOSLP model solution may not cover the impacted OD pairs and correct the demand estimation errors of the associated OD pairs. For this reason, a stochastic optimal sensor location problem (SOSLP) model is developed, based on the DOSLP model, by incorporating the network uncertainty into the model formulation.

6.1.1 Deterministic Optimal Sensor Location Problem (DOSLP)

The ability to observe flow patterns and performance characteristics of dynamic transportation systems remains an important challenge for transportation agencies, notwithstanding continuing advances in surveillance and communication technologies. In order to improve the efficiency of data collection and data support to the new generation of real-time network traffic estimation and prediction systems, it is critical to understand how sensor deployment affects the network observability and the estimation quality of network states.

This dissertation presents the deterministic optimal sensor location problem (DOSLP) under two situations, without and with budgetary constraints. In the first situation, the senor location problem is viewed as an O-D covering problem based on dynamic traffic assignment methodology. It is formulated as a binary integer programming model (DOSLP-1). However, in most real world applications, the number of sensors is constrained by budget/resource limitations. Aware of the inherent connection between the OD estimation problem and the sensor location problem, a Kalman filtering based model (DOSLP-2) is presented to explore time-dependent maximal information gains and O-D demand coverage across all links in the network in the second situation.

The Branch-and-bound (BnB) method, which is commonly used to solve computationally intensive integer problems, is used to solve DOSLP-1. Recognizing that the DOSLP-2 model is non-convex, the solution procedure is formulated as a bi-level stochastic integer program. The upper-level seeks potential locations according to some selection rules, while at the lower level, the selected locations are evaluated using the simulated results by running a user equilibrium simulation-based DTA procedure (in this case using the DYNASMART-P softwar). A hybrid greedy randomized adaptive search heuristic is developed for efficiently exploiting the near optimal sensor locations because of the network size and computational complexity of the proposed problem.

6.1.2 Stochastic Optimal Sensor Location Problem (SOSLP)

Uncertainty is one of the major factors that transportation system analysts and planners have to deal with in making transportation planning decisions. It plays a critical role since transportation agencies and planners have to deploy limited sensors in the network before the occurrence of unanticipated events (e.g. incidents, weather, special events, etc), which will subsequently impact the vehicle paths and traffic pattern in the network. Thus, the quality of the network state and estimated trip matrix may be impaired because the sensor location solutions from DOSLP may not be able to capture the impacted OD flows in the occurrence of the uncertain events. Based on a two stage stochastic model and iterative bi-level solution framework, this research extends the DOSLP to a stochastic problem and proposes a robust formulation SOSOLP to accommodate the un-anticipated network events in seeking to achieve the objectives of enhancing the long-run expectation of OD demand estimation quality and maximizing the long-run expectation of OD flow coverage under stochastic network environments.

By assuming that the occurrence of incidents on a link follows a Poisson process, and that likelihood of incident occurrence on a link is obtained from Bayesian statistical method, a modified HGRASP-DTA search procedure of DOSLP is used to find the near optimal sensor locations based on dynamic traffic assignment methodology.

6.1.3 Research Findings

To circumvent the difficulties of obtaining real-time link count data and historical variance and covariance of the OD demand, this research uses a synthetic data set from a DTA-based simulator, DYNASMART-P, albeit for a real network configuration, to evaluate the performance of the proposed models. The extensive numerical experiments conducted as part of this research resulted in the following key findings:

1. Sensors should be located on those links so that they can maximally intercept OD flows;
2. The sensor observation data should be linearly independent;
3. Adding more sensors does not always generate larger information gain;
4. The lower the measurement error is, the more gains the system could attain;
5. Maximization of sensor network coverage does not necessarily yield the largest improvement in the overall OD demand estimation quality;
6. In the presence of network uncertainty, a two-stage stochastic model accounting for impacted vehicles can provide more robust and accurate estimates than the deterministic model for OD demand flow and network link performance.
7. The sensor location strategies from the proposed models provide more robust and accurate demand estimates and larger OD flow coverage than the existing static sensor location model based on the static traffic assignment methodology.

6.2 Research Contributions

This section presents specific contributions of this research to the theoretical and algorithmic development of the sensor location problem in a large-scale traffic network.

To date, there are few studies conducted on the sensor location problem. Most of the existing methodologies for sensor networks are based on static traffic assignment assumptions and use OD flow coverage and flow capture as the objectives to locate the sensors. The major limitation of the static sensor location models is that they cannot capture the traffic interaction among vehicles and adjacent links, which may result in solutions that could perform worse than placement based on general engineering judgment. Another drawback of the static sensor location models is their inability to capture the traffic dynamics, such as the vehicle path evolution, especially in a congestion network. Furthermore, an important finding in this study is that locating sensors exclusively on the basis of flow coverage maximization does not necessarily lead to the largest improvement in the overall OD demand estimation quality. In addition, existing static sensor location models either lack efficient solution procedure for actual large-scale networks, or are unable to respond to the network uncertainty and its consequential impact on traffic conditions.

To circumvent the principal difficulties in estimating the dynamic link proportion matrices, a dynamic traffic simulator is used to propagate the vehicles along the user equilibrium paths and determine the system state. Based on the simulation-based solution methodology for dynamic traffic assignment (Mahmassani 1998), this dissertation provides dynamic models that can be used to gain insight into the sensor location problem when traffic dynamics and network uncertainty are accounted for. It generalizes the static traffic assignment assumption and exploits the optimal sensor location strategies based on dynamic traffic assignment methodology. With a heavy emphasis on the OD demand estimation problem under different scenarios, Kalman filtering based dynamic sensor location model formulations and their associated algorithms are constructed to find robust solutions for real-time estimation and prediction applications in large-scale networks. This dissertation provides the following key contributions to the sensor location problem:

- This research introduced a new perspective in the sensor location problem. It interpreted the sensor location problem as a value of information problem, which leads the problem to the learning process models. The proposed Kalman filtering model based bi-objective framework provides a flexible and tractable approach to incorporate OD flow coverage and demand uncertainty reduction. In addition, it introduced the traffic dynamics and network uncertainties into the sensor location problem formulation, which essentially captures the impacted OD flows when the uncertainty unfolds.
- This research explored ways to allocate resources to create a network detection system in a manner that produces minimal estimation errors and minimal costs. It systematically analyzed the relationships among the sensor locations, time-dependent OD coverage and demand estimation error correction. The sensitivity analysis on the effects of sensor locations and sensor numbers to the network status estimation and prediction reveals the importance of optimal sensor locations in a large-scale network.

- A two-stage stochastic model provides an integrated framework to account for the inherent uncertainty in traffic networks in the sensor location problem. It proposed an incident generation model and considered multiple incident scenarios. Furthermore, the proposed SOSLP model classified the impacted and un-impacted vehicles into different classes to minimize the user travel time by diverting the impacted vehicles to alternative routes.
- This research proposed an effective and tractable solution procedure, the Hybrid Greedy Randomized Adaptive Search Procedure (HGRASP-DTA). The procedure is used to solve large-scale *NP-hard* problem for general traffic networks. It efficiently searches for robust solutions in the feasible domain for the sensor location problem in the context of real-time traffic estimation and prediction in large scale networks

In sum, this dissertation systematically proposes a new methodology to exploit the robust sensor locations for the purpose of real-time traffic estimation and prediction to support advanced traffic management and traveler information systems in an urban transportation network context. More importantly, this dissertation strengthens the inherent connection between sensor location and the demand estimation problem, rather than formulating this problem as an OD coverage problem under static traffic assignment. The proposed models and solution procedures were systematically integrated into off-line and on-line DTA systems, and were rigorously tested and evaluated using field data as well as synthetic data based on several realistic networks.

6.3 Future Research and Extensions

As an initial effort in introducing traffic dynamics and incorporating network uncertainty into the sensor location problem, several aspects of the proposed framework and solution algorithms in this dissertation leave room for further investigation and improvement. This section outlines several major directions of future research and extensions of this dissertation.

(1). Extension to other network state estimation and prediction applications

In this research, the proposed models underline a methodology to optimally deploy limited sensors for demand estimation error correction. The first natural extension is to incorporate new observed data sources into the model formulation and find the optimal locations for those new observation facilities as a supplement to the traditional point sensors. The second extension for future research is to deploy the sensors for other traffic state estimation and prediction applications, such as estimated travel time reliability, measuring and predicting traffic travel time and link state estimation, etc. Furthermore, it is useful to find second-optimal sensor locations in addition of the existing sensors for a sensor network. Another extension of the sensor location problem is to develop an effective framework for integrating the OD decomposition strategy into the online simulation-based DTA system.

(2). Develop the sensor location problem as a chance-constrained model

The proposed two-stage stochastic optimal sensor location model under uncertainty provides new insight for deploying sensors in a realistic large scale network. A further study would be to formulate the sensor location problem as a chance-constrained model to maximize the probabilities of reaching certain goal, such as introducing a stochastic threshold constraint. The constraint could be that a link will be considered to install a sensor only if a minimum level of demand is captured or a minimum level of link information gain is obtained at that site.

(3). Evaluate the proposed dynamic sensor location model with real-world data and more diverse types of data sources.

In order to circumvent the difficulties of obtaining the real-world link count data, historical variance-covariance data and the incident occurrence probabilities, the proposed dynamic sensor location models were investigated using synthetic data sets. However, the real-world data contain actual traffic information and can provide further insight on the effects of actual data to the performance of models. For example, network simultaneously occurred incidents are assumed independent with each other and synthetic incident severity was used in this research, the actual observation data could be used to further evaluate the robustness of the proposed models under the realistic traffic conditions. In addition, using data from other sources could be considered to increase the network observability and enhance the demand estimation quality. For example, AVI (automated vehicle identification) data, or link densities obtained from processed video imaging data, could be incorporated into the proposed location model.

(4). Develop a more efficient and tractable solution algorithm for large-scale networks

Since the properties and general efficient method to solve large scale stochastic integer programming are scarce, efficient algorithms and solution procedures, such as decomposition methodologies to solve the proposed two-stage stochastic model are undoubtedly needed.

Bibliography

Aitken, A.C., 1935. On least squares and linear combinations of observations. Proceedings of the Royal Statistical Society, 55, 42-48.

Ashok, K and Ben-Akiva, M. E (1993)., Dynamic O–D Matrix Estimation and Prediction for Real-Time Traffic Management Systems, in Transportation and Traffic Theory, C. F. Daganzo (ed), Elsevier, New York, 465–484.

Ashok, K and Ben-Akiva, M. E (2000). Alternative approaches for real-time estimation and prediction of time-dependent origin destination flows, Transportation Science, Vol 34, No 1 pp21-36

Ashok, K and Ben-Akiva, M. E (2002). Estimation and Prediction of Time-Dependent Origin-Destination Flows with a Stochastic Mapping to Path Flows and Link Flows, Transportation Science, Vol 36, No 2 pp184-198

Beale, E.M.L. (1955), On Minimizing a Convex Function Subject to Linear inequalities, Journal of Royal Statistical Society, series B17, pp173-184

Bell, M.G.H. 1991. The real-time estimation of origin-destination flows in the presence of platoon dispersion. Transportation Research. 25B, pp115-125

Bell, M.G.H. 1991. The estimation of origin-destination matrices by constrained generalized least squares. Transportation Research 25B, 13-22

Ben-Akiva, M. E. and Lerman, S. R. (1985). Discrete Choice Analysis: Theory and Application to Travel Demand, MIT Press, Cambridge, Ma.

Bianco, L., Confessore G. and Reverberi, P. (2001). A Network Based Model for Traffic Sensor Location with Implications on O–D Matrix Estimates, Transportation Science 35, pp. 50–60.

Bierlaire, M. and Crittin, F. (2004). An efficient algorithm for real-time estimation and prediction of dynamic OD tables, Operation Research Vol 52, No.1, pp116-127.

Birge, J.R., and Louveaux , F. (1997), Introduction to Stochastic Programming. Springer-Verlag New York, New York, NY.

Cascetta, E. 1984. Estimation of trip matrices from traffic counts and survey data: A generalized least squares estimator. Transportation Research, 18B, 289-299.

Cascettae, E and Cantarellag, G (1991), A Day-to-Day and Within-day Dynamic Stochastic Assignment Model, Transportation Research, 25A

Cascetta, E., Inaudi, D. and Marquis, G. (1993). Dynamic estimator of Origin-Destination matrices using traffic counts. Transportation Science, Vol.27, No.4, 1993, pp363-373

Center for Advanced Transportation Technology, Research Center at the University of Maryland, College Park, MD. July 2004.
http://www.cattlab.umd.edu/cf/index.cfm?js=enabled&bin=trafficData

Chootinan, P, Chen, A and Yang, H. 2005. A bi-objective counting location problem for origin-destination trip estimation. Transportmatrica Vol.1. No. 1. 65-80

Chiu, Y. C.,Huynh,N and Mahmassani, H. S. (2001), Determining Optimal Locations for Variable Message Signs Under Stochastic Incident Scenarios, CD-ROM paper Preprints, The 80st Annual Meeting of TRB, January, Washington D.C., USA.

Chiu, Y. -C., Mahmassani, H. S. (2002), Hybrid Real-Time Dynamic Traffic Assignment Approach for Robust Network Performance, CD-ROM paper Preprints, The 81st Annual Meeting of TRB, January, Washington D.C., USA.

Chui, C.K. and G. Chen, 1991, Kalman filtering with real-time applications, IEEE Transactions on Aerospace and Electronic Systems, 27(1), 149-154.

Cremer M. and H. Keller. Dynamic identification of flows from traffic counts at complex intersections. Proc. 8th Int. Symp. Transportation and Traffic Theory, Toronto University, 1981.

Dantzig, G.B. (1955), Linear Programming under Uncertainty, Management Science. Vol.1, pp.197-206

Dixon, Michael P. and Rilett, L.R. 2002. Real-Time OD estimation using automatic vehicle identification and traffic count data. Computer-Aided Civil and Infrastructure Engineering. 17, 7-21

Doan, D. L., Ziliaskopoulos, A. and Mahmassani, H.S. (1998). On-line Monitoring System For Real-time Traffic Management Applications, Presented at the 78th Annual meeting of the Transportation Research Board, Washington D.C

Ehlert, A., M.G.H. Bell, and S. Grosso, 2006. The optimization of traffic count locations in road networks, Transportation Research Part B, 40, 460-479.

Eisenman, S.M., Fei, X., Zhou, X.S and Mahmassani, H.S, 2006. Number and location of sensors for real-time network traffic estimation and prediction: A sensitivity analysis. Transportation Research Record 1964, pp253-260.

Fei, X., Eisenman, S.M and Mahmassani, H.S., 2007. Sensor Coverage and Location for Real-time Traffic Prediction in Large-Scale Networks. 86th Transportation Research Annual Meeting CD, Washington DC.

Feo, T.A. and Resende, M.G..C. 1995. Greedy randomized adaptive search procedures, Journal. of Global Optimization, vol. 6, 109-133

Festa, P. and Resende, M.G..C. 2001. GRASP: An Annotated Bibliography, AT&T Labs Research Technical Report.

Fisk, C.S. 1988. On combining maximum entropy trip matrix estimation with user optimal assignment, Transportation Research, 22B, 69-79

Gelb, A., 1974. Applied Optimal Estimation, MTI Press, Cambridge, MA.

Gendreau ,M., Laporte ,G., and Seguin, R.(1996), Stochastic Vehicle Routing. European Journal of Operational Research, 88, pp3–12.

Gentili, M and Mirchandani, P.B. 2005. Locating Active Sensors on traffic networks, Annals of Operations Research 136, pp297-257.

Geoffrion, A.M.(1970), Elements of Large-Scale Mathematical Programming: Part I: Concepts, Management Science, Vol. 16. No.11, Theory Series, pp652-675

Glover, F. (1987), Tabu Search Methods in Artificial Intelligence and Operations Research, ORSA Artificial Intelligence 1, No. 2.

Glover, F., and Laguna, M., (1993). Tabu search. In: Reeves, C. R. (Ed.), Modern Heuristic techniques for Combinatorial Problems. Blackwell Scientific Publications, Oxford.

Greene, W.H., 2000. Econometric Analysis, 5th ed., Prentice Hall, Upper Saddle River, NJ.

Hawas, Y. and Mahmassani, H.S., "Comparative Analysis of the Robustness of Centralized and Distributed Route Control Systems in Incident Situations," Transportation Research Record. 1537, pp. 83-90, 1997.

Jiang, Hai (2004), Parallel implementations of dynamic traffic assignment models and algorithms for dynamic shortest path problems, Massachusetts Institute of Technology. Dept. of Civil and Environmental Engineering. Master Thesis, 2004

Kachroo, P., Narayanan, A. and Ozbay, K., 1995. "Investigating the Use of Kalman Filtering Approaches for Dynamic Origin–Destination Trip Table Estimation," Working paper, Center for Transportation Research, Virginia Tech, Blacksburg, VA, 1995.

Kang, Ying., 1999 Estimation and prediction of dynamic Origin-Destination(O-D) demand and system consistency control for real-time dynamic traffic assignment operation. The University of Texas at Austin, May 1999

Kirkpatrick, S., Gelat, C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. Science 220, 671–680.

Lam, W. H. K., and H. P. Lo, 1990. Accuracy of O-D estimates from traffic counting stations. Traffic Engineering and Control, 31, 358-367.

Lin, S. and Kernighan, B.W. 1973. An effective heuristic algorithm for the traveling salesman
problem. Operations Research, 21, 498-516.

Liu, Henry X., Ma, Wenteng, Jayakrishnan,R, Recker, Will (2004), Large-Scale Traffi c Simulation Through Distributed Computing of Paramics, California Path Research Report, UCB-ITS-PRR-2004-42, University of California, Berkeley.

Liu, Changzheng and Fan, Yueyue (2007), A Two-Stage Stochastic Programming Model for Transportation Network Retrofit, Transportation Research Board 86th Annual Meeting CD-ROM, Transportation Research Board, Washington DC.

Liu, Y, Lin, P.W., Zou, N, Chang, G.L., and Point-Du-Jour, J. Y. (2004), Emergency Incident Management, Benefits and Operational Issues – a Performance and Benefits Evaluation of CHART, The 2004 IEEE International Conference on Networking, Sensing and Control, Taipei, Taiwan, March 2004.

Lo, H.P., Zhang, N. and Lam, W.H.K. (1999) Decomposition algorithm for statistical estimation of OD matrix with random link choice proportions from traffic counts, Transportation Research Part B, No.33, 369-385

Maher, M.J. 1983. Inferences on trip matrices from observations on link volumes: A Bayesian statistical approach. Transportation Research, 17B, 435-447.

Mahfoud, R. H. (2002) Distributed Software Design and Implementation of a Real-Time Traffic Estimation and Prediction System, M.S. Thesis, The University of Texas at Austin.

Mahmassani, H (1984), Uncertainty in transportation systems evaluation: Issues and approaches. Transportation Planning and Technology, Vol. 9, pp. 1-12.

Mahmassani, H.S. (1998). "Dynamic Traffic Simulation and Assignment Models, Algorithms and Applications to ATIS/ATMS Evaluation and Operation", in Operations Research and Decision Aid Methodologies in Traffic and Engineering Management, Edited by Toint, Labbe, Laporte and Tanczos, Springer-Verlag, Berlin, 104-135.

Mahmassani, H. S., Hawas, Y., Abdelghani, K., Abdelfatah, A., Chiu, Y-C. and Kang, Y. (1998) DYNASMART-X; Volume II: Analytical and Algorithmic Aspects. Technical Report ST067-85-Volume II, Center for Transportation Research, The University of Texas at Austin.

Mahmassani, H.S., Abdelghany, A., Chiu, Y.-C., Huynh, N., Zhou, X. 2000. DYNASMART-P, Intelligent Transportation Network Planning Tool Volume I, II. Report ST067-85-PII, prepared for Federal Highway Administration, January, Austin, Texas, USA. University of Texas at Austin, Austin.

Mahmassani, H.S., A.F. Abdelghany N. Huynh, X. Zhou, Y-C. Chiu, and K.F. Abdelghany. DYNASMART-P (version 0.926) User's Guide. Technical Report STO67-85-PIII, Center for Transportation Research, University of Texas at Austin. 2001.

Mahmassani, H.S., and Zhou, X., 2005. "Transportation System Intelligence: Performance Measurement and Real-Time Traffic Estimation and Prediction in a Day-to-Day Learning Framework, " Chapter 16 in Advances in Control, Communication Networks, and Transportation Systems, In Honor of Pravin Varaiya, edited by E. Abed, Birkhauser

Mahmassani, H.S., Zhou, X and. Qin, X, (2002). TREPS Phase 1.5B Final Report: DYNASMARTX Evaluation for Real-Time TMC Application: Irvine Test Bed. Prepared for the U.S. Department of Transportation,Maryland Transportation Initiative

Martin, P.T, Perrin,J and Hansen, B. (2001), Incident Detection Algorithm Evaluation, University of Utah, Prepared for Utah Department of Transportation, March 2001

Maryland Manual on Uniform Traffic Control Devices for Streets and Highways (2006 Edition), Part 6, Temporary Traffic Control, Maryland State Highway Administration.

Mulvey, J.M., Vanderbei , R.J. and Zenios , S.A (1995), Robust Optimization of Large-Scale systems Operations Research, 43(2), 264-281.

Okutani, I. (1987), “The Kalman Filtering Approach in Some Transportation and Traffic Problems,” In Transportation and Traffic Theory, N. H. Gartner and N. H. M. Wilson (eds), Elsevier, New York, 397–416.

Patriksson, M. and Wynter, L.(1999), Stochastic Mathematical Programs with Equilibrium Constraints. Operations Research Letters, 25, pp159-167.

Peeta S. And Mahmassani, H.S. 1995. System optimal and user equilibrium time-dependent traffic assignment in congested networks, Annals of Operations Research, 60, pp 81-113.

Peeta, S. and Chen, S-C. (1999). "A Distributed Computing Environment for Dynamic Traffic Operations," Computer-Aided Civil and Infrastructure Engineering, Vol. 14, pp. 257-271.

Peeta, S. and Zhang, P. (2004). "On-Line Control Architecture for Enabling Real-time Traffic System Operations," Computer-Aided Civil and Infrastructure Engineering; Vol. 19, pp. 306-323.

Peters,M.E, (2007), National Strategy to Reduce Congestion on America’s Transportation Network, www.fightgridlocknow.gov.

Pitsoulis, L. S. and Resende, M.G..C. 2001, Greedy randomized adaptive search procedures, AT&T Labs Research Technical Report.

Pravinvongvuth, S., Chen, A., Chootinan, P. and Narupiti,S. 2005. A methodology for selecting pareto optimal solutions developed by a multi-objective AVI reader location model. Journal of the Eastern Asia Society for Transportation Studies. Vol 6 2441-2456

Sawaya, O B., Doan, D. L., Ziliaskopoulos, A. K. and Fourer, R (2001), A Multistage Stochastic System Optimum Dynamic Traffic Assignment Program with Recourse for Incident Traffic Management, Transportation Research Record 1748, pp 116-124

Scales, L.E. 1985. Introduction to Non-linear optimization. MacMillan, New York.

Schrank, D. L and Lomax, T. J. (2005) The 2005 Urban Mobility Report, Texas Transportation Institute, Texas A&M University, College Station, TX.

Sen, S and Higle, J.L., (1999), An introductory tutorial on stochastic linear programming models, Interfaces 29, pp33-61

Spiess, H. 1987. A maximum-liklihood model for estimating origin-destination matrices, Transportation Research 21B, 395-412

Tavana, H and Mahmassani, H.S 2000, Estimation of dynamic origin-destination flows from sensor data using bi-level optimization method, Preprint paper No# 01-3241,Transportation Research Board, DC.

Waller, S.T., J.L. Schofer, and Ziliaskopoulos, A.K. (2001), Evaluation with Traffic Assignment under Demand Uncertainty, Transportation Research Board 1771, pp. 69-75.

Waller, S.T.and Ziliaskopoulos, A.K. (2001), Stochastic Dynamic Network Design Problem, Transportation Research Board No. 1771, pp. 106-113.

Wardrop, J.G. (1952). Some Theoretical Aspects of Road Traffic Research. Proceedings Institution of Civil Engineers II(1), pp. 325-378.

Wu, J.F and Chang, GL., 1996. Estimation of time-varying origin-destination distributions with dynamic screenline flows, Transportation Research 30B, No.4, pp277-290.

Wu,J.F. 1997. A real-time origin destination matrix updating algorithm for on-line applications, Transportation Research 31B, No.5, pp381-396

Yang, H., Y. Iida, and T. Sasaki. 1991. An analysis of the reliability of an origin-destination trip matrix estimated from traffic counts. Transportation Research, 25B, 351-363.

Yang H., and Zhou,J., 1998. Optimal traffic counting locations for origin-destination matrix estimation. Transportation Research B, 32(2): 109-126.

Yang, H, Yang, C. and Gan, L.P., 2006. Models and algorithms for the screen line-based traffic counting location problems. Computer & Operations Research 33. 836-858

Yim, K.N., and Lam, H.K. 1998. Evaluation of count location selection methods for estimation of OD matrices. Journal of Transportation Engineering, 124(4), 376-383

Yin, Y., Madanat, S. and Lu, X. (2005), Robust Improvement Schemes for Road Networks under Demand Uncertainty, 2005 TRB 84th Annual Meeting CD-ROM, Washington, D.C., January 2005.

Zhou, X.S. (2004). Dynamic Origin-Destination demand estimation and prediction for off-line and on-line dynamic traffic assignment operation. Ph.D. Dissertation, The University of Maryland at College Park.

Zhou, X.S. and Mahmassani, H.S. 2006. Dynamic Origin-Destination demand estimation using automatic vehicle identification data. IEEEE Transactions on Intelligent Transportation Systems. Vol 7. No.1, 105-114

Zhou, X.S. and List, G.F., 2006.An information-theoretic sensor location model for traffic origin destination demand estimation applications. Submitted for publication in Transportation Science

Zhou, X.S., Qin,X. and Mahmassani, H.S.(2003). Dynamic orgin-destination demand estimation with multiday link traffic counts for planning applications. Transportation Research Record 1831. pp30-38

Ziliaskopoulos,A, Kotzinos,D, and Mahmassani, H.S.(1997) Design and implementation of parallel time-dependent least time path algorithms for intelligent transportation systems, Transportation Research. 5C (2), pp. 95–107

Van Zuylen, H.J., and Willumsen, L.G.. 1980. The most likely trip matrix estimated from traffic counts. *Transportation Research Part B*, 14, 281-293.