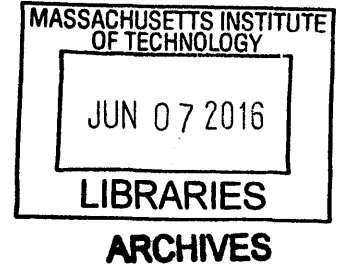# Real Time Toll Optimization based on Predicted Traffic Conditions

by

Shi Wang

Bachelor of Civil Engineering

Tsinghua University, Beijing, P.R. China (2010)

Submitted to the Department of Civil and Environmental Engineering

in partial fulfillment of the requirements for the degree of

Master of Science in Transportation

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2016

Signature redacted

Author . . . .

Department of Civil and Environmental Engineering

May 18, 2016

Signature redacted

Certified by . . . .

Moshe E. Ben-Akiva

Edmund K. Turner Professor of Civil and Environmental Engineering

Thesis Supervisor

Signature redacted

Accepted by . . . .

Heidi Nepf

Donald and Martha Harleman Professor of Civil and Environmental Engineering

Chair, Graduate Program Committee

Real Time Toll Optimization based on Predicted Traffic Conditions

by

Shi Wang

Submitted to the Department of Civil and Environmental Engineering

on May 18, 2016, in partial fulfillment of the

requirements for the degree of

Master of Science in Transportation

ABSTRACT

Road pricing is an effective method of demand management. Pricing on highway managed lanes is usually implemented as time-of-day or dynamic tolling in practice. Toll rates are usually updated according to latest traffic measurement and based on pre-defined rules. Researches on highway pricing can be generally categorized as analytical, reactive or optimization-based approaches. The limitations of current studies are compared and discussed in this thesis. A new framework is proposed which aims to develop an adaptive integrated simulation-optimization framework that brings together several enhancements: real time, predictive, simulation-based and consistent.

The main components of the framework include DTA model, DynaMIT, for evaluating control strategies, optimization module solving for optimal solution and real-life traffic system providing surveillance data. Optimization problem is formulated with rolling horizon scheme, and presented with basic models for revenue maximization. Close-loop testing approach is proposed by replacing traffic system with a microscopic simulator, MITSIM.

Tests are first conducted on a two-path synthetic network to demonstrate the capability of the framework with changing demand and different behavior parameters. Then a case study is performed on NTE Express Lanes network in Texas. Calibration of the network with multiple sources of traffic data is discussed, and initial calibration results with sensor data are presented. Also, the models are extended to account for the regulation rules imposed by the local government. Optimization results for morning peak period on a typical weekday are presented, and the resulting revenue is compared with the benchmark case. Finally, potential improvement in solution algorithm is discussed for the system's real time computational requirements.

The main contribution of the thesis includes: 1) identifying the limitations of

tolling strategies in practice and in academic researches, 2) proposing an adaptive integrated simulation-optimization framework, 3) demonstrating the capability of the framework through close-loop testing on a synthetic network, and 4) applying the framework on a real-world network with managed lanes, and proposing calibration approach incorporating multi-source traffic data.

Thesis Supervisor: Moshe E. Ben-Akiva
Title: Edmund K. Turner Professor of Civil and Environmental Engineering

Acknowledgement

Firstly, I would like express my sincere gratitude to my research advisor Professor Moshe Ben-Akiva for his guidance through these two years. Your guidance, expertise and patience supported my journey at MIT and contributed greatly to this thesis. Also, by taking your transportation economics and demand modelling courses, I learned a lot about how to approach complicated transportation problems. During discussion on research project, you always offer critical ideas and innovative insights on the results, and all these motivates me to think about better ways of dealing with new challenges.

I also want to thank Dr. Bilge Atasoy for your help during my work as research assistant. It is my great pleasure to collaborate with you on the research project. Your patience and insights are always the key to solving various problems during my research and delivering our progress to clients from CINTRA. Also, I want to thank you for the useful advice on my study and research at MIT.

Many others in the ITS lab and in the SMART team at Singapore helped me a lot during the past two years: Haizheng Zhang, Xiang Song, Yundi Zhang, Weikun Hu, Ravi Seshadri, Dong Wang, Samarth Gupta, Yang Lu. Your sincere suggestion and support on my study and research helped me go through various challenges at MIT. Also I would like to thank Eunice Kim and Katherine Rosa for your encouragement and help during my study at MIT. Thank you to CINTRA for sponsoring me with the research project, and to the Singapore-MIT Alliance for Research and Technology (SMART) for the summer internship at the end of my first year.

I would like to thank our collaborators from CINTRA, Ricardo Sanchez, Thu Hoang, Andres De Los Rios, Megan Rhodes, John Brady and Pablo Ferrando, for the continuous help and valuable feedback throughout the project.

Finally, I would like to thank my family, especially my father Yimin Wang and my mother Li Wang for your love and support, which gives me the power and confidence to overcome any difficulties in my life.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Overview of road pricing

Traffic congestion affects economic productivity as well as the environment and traffic safety. In order to ameliorate congestion, actions can be taken on the supply and demand sides. On the supply side, capacity can be increased through the construction of new infrastructures. However this method is often too costly and inefficient. An alternative approach of increasing interest is to improve the efficiency of transportation networks through demand management.

Road pricing is an effective method of demand management. The purpose of road pricing may vary by the road operators. Besides the most common objective to relieve traffic congestion, it can also be deployed to internalize the cost of vehicle emission, noise and road damage (de Palma and Lindsey, 2011). Private sectors may also use pricing as a tool to generate revenue. The discussion on road pricing covers a variety of aspects, including use of road pricing revenues, parking pricing, pricing of road emission, etc. Road pricing strategies can also be categorized in various ways like toll collecting technology, the degree to which the toll rates change by time, the definition of area where vehicles are subject to toll, etc. In this section, the main focus would be the categorization in spatial and temporal dimensions:

Spatially, congestion pricing may be classified according to the area of influence.

Facility-based scheme: Implementations include bridges, tunnels, managed lanes, etc. Tolls can be collected either at a single point or multiple points depending on distance travelled.

Cordons: A form of area-based charging where vehicles have to pay a toll when crossing the cordon in the inbound/outbound direction. Examples include EcoPass

introduced in Milan in 2008 to mitigate vehicle emission[1] and the cordon surrounding the city center of Stockholm to relieve congestion[2].

Zonal scheme: Vehicles have to pay a toll to enter/exit a zone, and the zone boundaries can be defined wither by natural features or by elements of the built environment. One notable example is the London congestion pricing introduced in 2003[3].

Figure 1.1 Singapore ERP system[4]
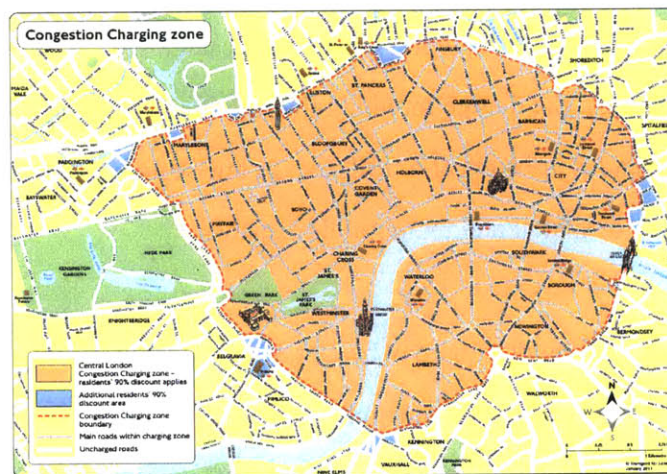
Figure 1.2 London congestion pricing[5]

---

[1] http://urbanaccessregulations.eu/ecopass

[2] https://www.epass24.com/

[3] https://tfl.gov.uk/modes/driving/congestion-charge

[4] http://www.lta.gov.sg/content/ltaweb/en.html

[5] https://tfl.gov.uk/modes/driving/congestion-charge/congestion-charge-zone

Distance-based scheme: Toll rates vary depending on the distance of the vehicle travelling on this facility, and the relationship between distance and toll rate can be either linear or non-linear.

Temporally, tolls are applied using two strategies: **fixed** and **dynamic**.

In a fixed pricing setting, the toll either stays flat or varies in a predetermined fashion during the day, which is often referred to as time-of-day tolling. One main reason for the prevalent use of fixed tolling lies in its straightforward setting of toll level and easy real time control. A fixed price setting took place in Oslo, Norway, where the congestion was reduced by 5% in the first year of the implementation of an urban toll ring project (Ieromonachou et al., 2006). The SR-91 Express lanes in Orange City (Yildirim, 2001) uses a time-of-day rate to accommodate for different demand levels across the day.

One main drawback of flat pricing is that it fails to consider the variability of price elasticity of demand over time (e.g., during peak and off-peak period) and the change in traffic network state. With the more advanced electronic toll collection technology, and the growing ease of change the toll rates in real time, flat pricing is now less preferable than time-of-day tolling or dynamic tolling.

Time-of-day tolling, as different from flat tolling, varies the toll rates by a given toll step (e.g., 1 hour, 30min) and the tolling schedules are adjusted periodically (e.g., monthly, seasonally) to account for the latest traffic surveillance information. It is more flexible compared with flat tolling.

Dynamic tolling in practice is often implemented in the form of responsive tolling, in which the latest toll rate is updated as a function of the prevailing traffic conditions. In real world applications, dynamic tolling has demonstrated traffic congestion alleviation and traffic safety improvement.

1.2 Pricing on managed lane

"Managed lanes" are defined by Federal Highway Administration in United States as "highway facilities or a set of lanes where operational strategies are proactively implemented and managed in response to changing conditions" (FHWA, 2008). Examples of managed lane include high-occupancy vehicle (HOV) lanes, value priced lanes, high-occupancy toll (HOT) lanes, or exclusive or special use lanes. There are various operational strategies applied to managed lane, like pricing, vehicle eligibility, access control, etc.

HOT lane (High Occupancy Toll lane) is special type of lane which allows HOV (high occupancy vehicle) and other exempt vehicles to travel without any charge. Other vehicles could have access to HOV lane by paying a variable toll which is adjusted in response to demand. Usually for a managed lane scheme, drivers have the choice to choose between the parallel uncharged lane (which is also called general purpose lane, or GP lane) and the HOT lane. Drivers are motivated to choose HOT lane mainly because of expected less congestion and faster travel time than general purpose lane (Göçmen et al., 2015). Some of the managed lanes are designed to maximize throughput, while others may focus on maintaining free-flow condition or revenue maximization.

Examples of tolling on managed lane

San Diego I-15 is a typical highway system with dynamic pricing operation on HOT lane facilities[6]. The 20-mile I-15 Express Lanes project runs from SR 78 in Escondido to SR 163 in San Diego. Entrances and exits are constructed every two to three miles for travelers to move on and off the main lanes to the Express Lanes. While carpools, vanpools, and other exempt vehicles can use the Express Lanes free of charge, single-occupant vehicles can also access the Express Lanes with a

---

[6] http://511sd.com/fastrak511sd/I-15ExpLanes

FasTrakpass to electronically pay the toll. Compared with fixed monthly pricing, dynamic pricing offers a more customized facility use with response to traffic demand, and free-flow condition is usually maintained by the time-varying toll rate (Supernak et al., 2003).

Figure 1.3 Map of I-15[7]

Figure 1.4 Map of I394[8]

---

[7] http://511sd.com/fastrak511sd/FasTrakHome

[8] http://www.mnpass.org/
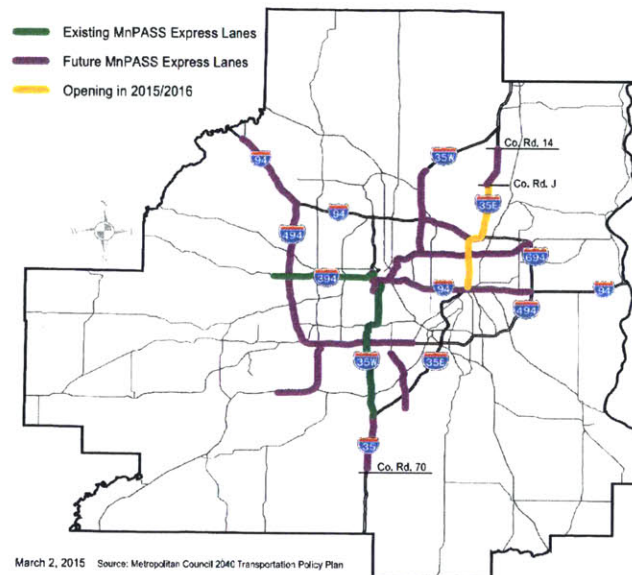
Minnesota's Mn/PASS I-394 HOT lane is another notable example of dynamic priced HOT lanes with multiple entries and exits[9]. The main goal of Mn/PASS is to improve the usage of the high occupancy vehicle (HOV) lane built in 1992, and the underused HOV lanes were converted to HOT lanes, also known as Mn/PASS 394 Express Lanes. Speeds and capacity of the HOT lane are maintained by dynamically changing the toll according to the demand and use of the lane. In-vehicle transponders, road scanners, and loop detectors are used to automate the processes of toll collection without stopping or slowing down drivers. Number of crashed was reduced by 5.3 percent based on the before-after data (Z. Xu and Huang, 2012), Overall, the efficiency of MnPASS Express Lanes has improved compared with pre-MnPASS condition as vehicle throughput has increased by 48% and person throughput by 25% (Buckeye, 2012).

Highway Tolling Algorithms

Tolls on highways represent an example of facility-based congestion pricing and mostly implemented as an operational policy on managed lanes, where strategies are developed in real time in response to changing conditions. The main purpose of developing toll lanes is to provide road users with congestion-free services and improve safety and operations through increased flexibility (Chung, 2013).

The tolling algorithms for highways, especially for time-of-day or dynamic tolling, also vary across the road operators. They may differ in the frequency of toll changing, measurement used to update tolls, lower and upper bound of toll rates, etc. Here are a few examples on how the operators determine the toll rates.

The SR-91 time-of-day toll schedule is updated seasonally, and the toll rates are decided by the directional average volume for each hour of the day during the past 12 consecutive weeks, and some if-then rules are used to calculate the toll rates for each

---

[9] http://www.mnpass.org/

hour (OCTA, 2003).

The I-15 South (of SR-56) develops a look up table to determine the toll rates for HOT lanes. 12-min volume lower thresholds, 6-min average volumes and level of service are used to decide the updated toll rate, and the main aim is to keep the HOT lane no worse than Level Of Service (LOS) C. There is a constraint on maximal of each toll adjustment every 6 minutes (SANDAG, 2006).

I-15 North (of SR-56) uses a more complicated approach which considers the different ingress/egress points, and divides the corridor into n zones with tolls derived from a series of steps, including value of time adjustment, travel time saving calculation and toll rate update. This procedure ensures that HOT lane is operating at free flow speed.

I-394 updates the toll rates every 3 minutes depending on the real time density and a pre-defined lookup table. The change in averaged density over the last six minutes is used to look up the incremental amount of toll values. Traffic on the general purpose lane is partly considered by the density change on HOT lane.

While each freeway operator adopts different type of tolling algorithms, there are some similarities.

(1) The traffic surveillance system will provide real time or averaged measurement of traffic speed/volume/density on the managed lane.

(2) Toll rate would increase in reaction to worsen condition on managed lane, and one important aim is to keep the free flow speed on managed lane.

(3) For dynamic tolls, toll rates are updated by a certain interval length to accommodate the traffic condition change.

(4) Drivers' reaction to tolls are not explicitly considered or oversimplified when deciding the algorithms.

(5) There is constraint on the toll incremental value during consecutive intervals, and there is upper bound specified by regulation.

1.3 Overview of DTA systems

Dynamic traffic assignment (DTA) system has the capability of capturing complex traffic dynamics for a large-scale traffic network. So it can be useful in providing prediction of traffic network state and evaluating traffic control strategies. The typical components of DTA models include demand models for travelers' trip decisions (OD, departure time choice, mode choice, route choice, etc) and supply models for traffic flow movement (Ben-Akiva et al., 2001; Florian et al., 2001).

Traffic simulators provide real time estimation and prediction of demand and traffic conditions, and they are of increasing interest given their valuable impact. Ramos et al. (2012) highlighted the need for modular platforms that integrate proper modeling and simulation capabilities in order to provide efficient management of operations. They work with AIMSUN[10], a microscopic traffic simulator. In urban freight transportation, Grzybowska and Barceló (2012) presented a decision support system for real time freight management with an integrated system of vehicle routing and dynamic traffic simulation models.

There are two main types of DTA system: **Analytical** and **Simulation-based**.

Analytical model is usually based on mathematical formulation and optimization. While the mathematical formulation gives computational advantage, the simplified assumptions usually degrades the model's capacity to capture the realistic traffic dynamics and travelers' trip decisions.

Simulation-based model uses more accurate traffic network representation, network state estimation and real time prediction. Based on the level of details, they can be further divided into microscopic, mesoscopic and macroscopic models. Microscopic and mesoscopic model have the capability of capturing drivers' travel

---

[10] http://www.tss-bcn.com

behavior and their real time response to traffic control/information, so they have been utilized to provide consistent traffic prediction and route guidance (Mahmassani, 2001; Y Wen et al., 2006; Yang Wen et al., 2008).

1.4 Thesis motivation

While dynamic road pricing is regarded as an effective way to improve the network efficiency and relieve traffic congestion, past practices or studies mainly implemented models which are based on simplified assumptions of traffic dynamics, drivers' route choice and network equilibrium. Also, most of the analytical models are tested only on small-scale synthetic networks, and the true effectiveness of the model are not validated against a real-life complex network.

Our approach is an adaptive integrated simulation-optimization framework that brings together several enhancements. First of all, the dynamic pricing methodology is a predictive one meaning that the toll is optimized based on predicted traffic conditions. A mesoscopic traffic simulator is integrated with the toll optimization in order to obtain state estimation and prediction. This simulator embeds several modules including demand simulation (OD estimation/prediction, travel time estimation/prediction, route choice models etc.), supply simulation, and online calibration. Consistent guidance generation is implemented in the prediction-based guidance process, so that the network conditions that are expected to be realized, given the anticipated users' reactions, coincide with those that form the basis for the travel guidance. The toll optimization module is in complete interaction with the simulator such that the optimized toll is decided with several iterations between the two rather than a single feedback function. Finally, the framework is developed with the aim of having a real time performance. This thesis focuses on developing the testing framework for the toll optimization problem, and implementing the new tolling model on a synthetic network and case study on the NTE TEXpress Lanes network in Texas.

1.5 Thesis outline

The remainder of the thesis is organized as follows. Chapter 2 presents a detailed literature review on different frameworks, models and cases of dynamic tolling problem. Chapter 3 proposes the formulation of the optimization problem which interacts with the traffic simulator and solves to obtain the optimal toll values. Chapter 4 presents the initial results of the framework on synthetic testing network to demonstrate its capability to react to real time demand change, accommodate the toll rates and meet the defined objectives. Chapter 5 presents a case study with the optimization results on the NTE TEXpress Lanes network. Finally, conclusions and future work are included in Chapter 6.

2 Literature review

2.1 State-of-practice tolling strategy

Fixed pricing lacks an interaction between changing traffic conditions on the network and motorists' willingness to pay. Conversely, in dynamic pricing the toll is updated based on demand prediction and varying traffic conditions. As examples of dynamic pricing applications, we refer to the I-15 lane in San Diego and I-394 high-occupancy toll (HOT) lane in Minnesota (Cambridge Systematics, 2006; Supernak et al., 2003; Yin and Lou, 2009). Base price for I-15 varies from $0.50 to $4.00 depending on the time-of-day and the price can be adjusted in real time in response to the traffic conditions. The toll price on I-394 varies from $0.25 to $4.00 and can be adjusted as often as every 3 minutes based on look-up tables. The FastLane project in Tel-Aviv adopts if-then rules for the adjustment of the toll with the objective of maximizing revenue and maintaining desired travel times (Bar-Gera, 2012). The idea is depicted in Figure 2.1. Recently, several new dynamic pricing projects have been initiated in the US, including the I-495 and I-95 Express Lanes in Virginia and the NTE TEXpress and LBJ TEXpress projects in the Dallas-Fort Worth area, where the toll is adjusted based on the observed congestion levels in the network.
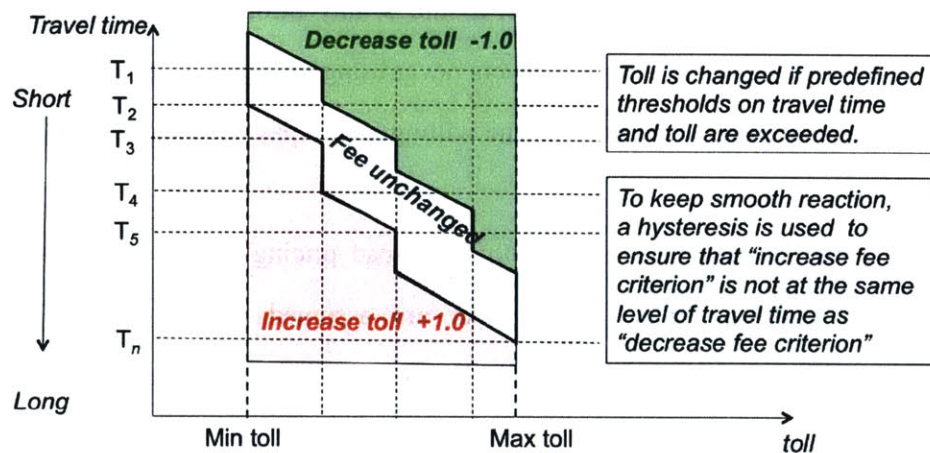
Figure 2.1 Illustration of rules used by Tel-Aviv

2.2 Reactive and predictive tolling strategy

State-of-the-practice for dynamic pricing mostly considers rule-based methods or look-up tables (Chung and Recker, 2011). More complicated algorithms have been studied in the literature. There are **reactive** vs **predictive** (also referred as *proactive* and *anticipatory*) methodologies.

Reactive dynamic pricing determines time-varying tolls in response to prevailing traffic conditions, and the idea has been utilized also for ramp metering (Papageorgiou et al., 1997). On the other hand predictive dynamic pricing methodologies incorporate the uncertainty in the system and dynamic pricing is carried out based on predictions. There are different conventions in terms of the way the uncertainties are taken into account: some consider scenario-based dynamic pricing and some others integrate dynamic pricing with traffic simulators in order to represent supply and demand as realistic as possible. Another distinction lies in the way the dynamic pricing decisions are applied to the real traffic system. Some of the studies consider an *offline* setting while others aim at having *real time* optimization for toll pricing, which is challenging in terms of computational time. In the remainder of this section we will give place to different studies that have different conventions in terms of the above mentioned concepts.

Li and Govind (2002) developed a decision model for assessing the tolling strategy for managed lanes. The model assumes fully informed traveler, fixed total demand and a certain proportion of drivers never using toll lane regardless of toll rate. The toll rate us jointly determined by the willingness to pay, time saving and flow on the two lanes.

Zhang and Levinson (2005) developed a road pricing model on a network of autonomous highway links. These links are assumed to be competitive and independent, and the objective is to maximize their own profit. Both decentralized autonomous links and centralized government control are considered to generate the pricing strategy.

Nagae and Akamatsu (2006) presented a case where a manager can switch between two levels of the toll in order to maximize the revenue taking into account the dynamic uncertainty of transportation demand. A dynamic programming formulation is adapted and an adjustment cost is considered for switching between the levels in order to avoid fluctuations.

Y. Wang et al. (2009) proposed a self-adaptive tolling algorithm for HOT lane toll optimization problem. Firstly, optimal flow ratio is determined based on observed traffic speed and toll changing pattern. Secondly, logit route choice model is used to calculate the proper toll rate to match the optimal flow ratio. Simulation-based evaluation in VISSIM for SR167 in Washington demonstrates the capability of the tolling algorithms in responding to real time traffic change.

Martin et al. (2009) presented a calibrated VISSIM model that integrates with HOT lane pricing. External travel route choice model is developed to capture drivers' route choice based on logit model. Travel time and traffic condition output is obtained from VISSIM simulation. Toll rates are updated to meet the predetermined proportion of vehicles that should be allowed on HOT lane.

Morgul (2010) employed traffic simulators Paramics and TransModeler for the evaluation of pricing algorithm. Toll rates are dynamically updated to be reactive to traffic conditions in a real time setting. The algorithm is of a rule-based type based on the observed occupancy and speed.

Jang et al. (2014) proposed a reactive pricing methodology based on the distribution function of value of time across the population. The toll rate is updated to react to the latest traffic condition. They present results with different objective functions including revenue maximization and delay minimization on a 14-mile freeway segment in the San Francisco Bay Area.

Boyles et al. (2015) incorporated departure time choice into the evaluation framework for HOT lane tolling. Drivers are divided into two groups; strategic drivers make their departure time choice so as to minimize expected generalized cost, and

choose either free lane or HOT lane, while captive drivers always choose free lane with random departure time. Various toll algorithms were compared under varying input demand.

Yin and Lou (2009) proposed two dynamic pricing approaches for managed toll lanes: *feedback-control approach* (Figure 2.2) and *reactive self-learning approach* (Figure 2.3). The pricing decisions are based on real time traffic conditions and the objective is to improve the free-flow travel service on the toll lanes while maximizing the freeway's throughput. A logit model is estimated in real time with a Kalman filtering technique, using the observed flow, travel time and toll rates. In an extended version of their model, they introduce heterogeneity in the logit model.
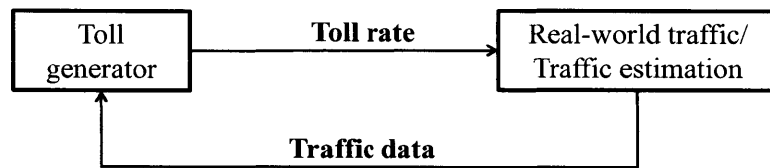
Figure 2.2 Feedback control
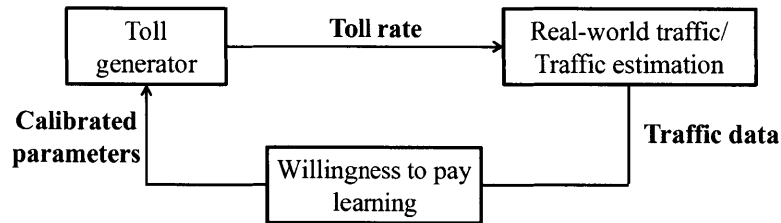
Figure 2.3 Reactive self-learning

Dong et al. (2011) proposed an anticipatory control for the pricing problem where the predicted traffic conditions are considered in order to generate the toll. As illustrated in Figure 2.4, the toll is determined by adjusting the previously predicted toll based on the deviation from the desired level. The dynamic pricing is integrated with DYNASMART, a real time traffic estimation and prediction system, which enables the modelling of users' response to price and road condition.

26
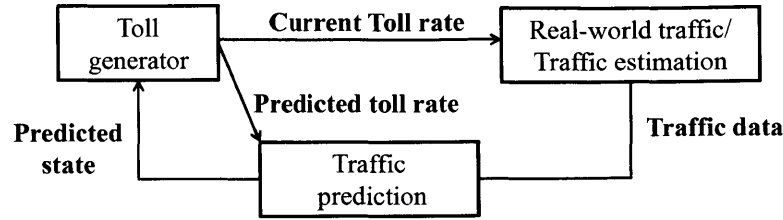
Figure 2.4 Anticipatory control

2.3 Optimization-based tolling strategy

Another approach, compared with reactive control, is to formulate the decision of road pricing strategy as optimization problem. Often there are two fundamental parts of such framework: formulation of performance objective and traffic modelling. The performance objective is often very flexible, and may vary by researches.

Joksimovic et al. (2005) formulated the pricing problem as a bi-level optimization problem, with the upper level model describing the network performance with given toll values, and the lower level model capturing traffic dynamics, route choice and departure time choice. Analytical form of DTA model is used to apply on two-path simple network. User heterogeneity is considered by assuming class-specific value of time.

Michalaka et al. (2011) developed a scenario-based toll optimization model in order to take into account the uncertainties in traffic demand. A two-stage stochastic formulation is adopted. The first stage determines the targeted inflows with the objective of maximizing the throughput of the freeway while ensuring a desirable density on the toll lane. The second stage optimizes the toll that would meet the targeted inflows. The scenario-based optimization minimizes the difference between the resulting and optimized demand on toll lane in high-consequence scenarios.

L. Yang et al. (2012) formulated a distance-based pricing problem on the basis of a stochastic macroscopic traffic flow model. Model assumes that switching traffic flow has no impact on the density evolution of free lane, and travel time on free lane

27

is not affected by tolling. A numerical algorithm is proposed for achieving optimal tolls in real time with the objective of maximizing total expected revenue.

Göçmen et al. (2015) developed the simulation model by three modules: demand generation, consumer choice and simulation for traffic dynamics. The traffic dynamics are modelled in the same way as the mesoscopic simulations used in DYNASMART. With the objective to be maximizing total revenue for the whole simulation period, different optimal policies are proposed to obtain the optimal time-dependent tolls.

Further researches use traffic simulators to capture both traffic dynamics and traveler behavior through the embedded models. Compared with analytical form of DTA model usually applied on simple testing network, simulation is more capable of capturing the complexity of large-scale network, and more applicable to modelling real-life network.

S. Xu (2009) reviews different pricing strategies and proposes a simulation-based optimization approach in an offline setting, where a toll optimization module sends the optimal toll to the real traffic system. Toll optimization module interacts with DynaMIT so that the dynamic pricing methodology has a predictive nature. The objective is to minimize the total travel time on the network assuming that total demand is fixed regardless of the toll rate.

Zou and Kulkarni (2013) developed a modelling and optimization framework to address the optimization problem for ramp metering and road pricing for High Occupancy Toll lane. Paramics is used as simulation model for traffic dynamics, and genetic algorithm based optimization engine is deployed to interact with traffic simulator and find optimal solution.

Toledo et al. (2015) developed a simulated-based control framework for real time toll optimization. Macroscopic traffic simulation model is used to describe traffic dynamics, and a logit model is used to predict drivers' route choice. Consistency between predicted travel time and information provided to users are achieved by

running iterations until convergence. Exhaustive search is used to solve the optimization problem

However, simulation for a long-scale network is often associated with high computation costs, and non-closed-form simulation model makes it hard to directly apply traditional optimization techniques. To address the high computational cost, surrogate-model approach, which is useful to solve a black-box optimization problem (Liem, 2007), has been proposed to approximate the true simulation function and reduce number of runs of simulation.

Surrogate-based optimization represents a class of optimization methodology, which utilizes surrogate modelling techniques to search for the optima. In such a framework, traditional optimization algorithms like gradient-based or evolutionary algorithms are only used for sub-optimization, while an upper-level surrogate model (or response surface model, meta-model) is used to improve the design efficiency and reduce number of runs of the high-fidelity, expensive simulation model (Han and Zhang, 2012).

X. Chen et al. (2015) proposed a simulation-based optimization algorithm to develop optimal traffic signal timing plan for a urban network, with a lower accuracy analytical model approximating the high accuracy simulator to improve computational performance. Trust-region algorithm is adapted for the optimization problem to make joint use of low-fidelity model and high-fidelity model in Amisun.

He et al. (2016) built a simulation based optimization framework for a time-varying pricing problem. To solve the multi-objective and constraint optimization problem, regression Kriging model, which is one typical type of function used for surrogate model, and infill strategies are used to approximate the true function. DynusT is used as the high-fidelity simulator to evaluate the input points.

2.4 Summary

The literatures can be generally categorized into three types: analytical approach, feedback control and optimization-based.

The analytical models usually have sound mathematical property, which makes it possible to study complex derivation of the model. The first-best pricing theory, for example, aims to set the pricing to be equal as the marginal cost of the road users. However, while these theories provide fundamental insights to the tolling problems, they can hardly be applied to real-world traffic system and highway tolling problems.

One main advantage of feedback control is its capability to react to real time changes in traffic conditions. Main challenges for feedback control lies in the control parameter tuning, which mainly relies on trial and errors in current research, and there is no standard for selecting the most suitable parameter value to achieve the predefined goal. Also, if the objective becomes more complicated than simply maintaining free-flow condition on managed lane, such as maximization of revenue or minimization of average travel time, feedback control is no longer a feasible approach.

Optimization-based approach offered the possibility of considering more complicated objective, and it is more flexible in including different constraints. While solving the optimization problem helps achieve the optimal control strategy, computation time and efficiency should be considered and well addressed if the optimization process in done in real time manner. If optimization is combined with simulation-based DTA model, the network complexity, traffic dynamics and users' travel behavior can be better captured.

3 Methodology

3.1 DynaMIT

In this research project, DynaMIT (Dynamic Network Assignment for the Management of Information to Travelers) is used to guide dynamic pricing decisions. DynaMIT is a mesoscopic simulator developed by the MIT ITS Lab that provides dynamic state estimation and prediction of traffic conditions. It combines models of travel demand and supply in order to estimate demand and traffic conditions and perform short-term predictions.

In real time applications, DynaMIT operates on a rolling horizon where estimation and prediction intervals are performed successively. The prediction horizon can be specified depending on the type of application. One main strength of the software is that the predictions provided by DynaMIT are consistent with traffic conditions that the users will face when they make their decisions (Bottom, 2000).

Demand

Disaggregate demand representation is employed in DynaMIT to model individual traveler's pre-trip and en-route decision. Individual's socioeconomic characteristics and access to real time information is considered during the decision process. Aggregate demand representation is used for time-dependent OD matrix. The historical OD flows are adjusted by DynaMIT in the following two stages (Ben-Akiva et al., 2010a):

(1) Drivers' response to real time information captured by disaggregate behavior model, which predicts travelers' departure time change, route choice and route switching.

(2) Online calibration of aggregated OD matrix based on real time traffic surveillance data.

Supply

The simulated network in DynaMIT consists of static and dynamic components. While the static part represents network topology, the dynamic part captures traffic dynamics. The main classes of models are employed in DynaMIT to capture the complexity of the network traffic flow:

(1) Capacity associated with road elements, incidents, controls.

(2) Deterministic queueing models for queueing part.

(3) Macroscopic speed density relationship for moving part.

The simulation of traffic operation is time based, and the process is divided into two phases: update phase and advance phase. The traffic dynamics like densities and speeds for road segments are updated during update phase, while the vehicles are advanced to new positions during the advance phase.

Route choice model

Route choice model in DynaMIT is categorized into pre-trip and en-route choice models. Here, we present the general structure with main features. It is formulated as a path size logit model, in which a path size variable is included for considering the overlapping of the routes (Bekhor et al., 2002). The utility function for path i is given as follows:

$$V_i = \mu GC_i + C_i \qquad 3.1$$

where GC_i is the generalized cost and C_i represents all other terms (e.g. early/late arrival, trip purpose, path size, etc) in the utility function including the path size. μ is the scale parameter that represents the price parameter since we have a generalized cost specification. GC_i is specified as follows:

$$GC_i = VOT \times TT_i + Toll_i \qquad 3.2$$

where TT_i is the travel time, VOT is the value of time for the given individual traveler, and $Toll_i$ is the toll value imposed on path i. VOT is assumed to be

32

randomly distributed across each traveler group according to a lognormal distribution.

Then is probability for the given traveler to choose path i is:

$$P_i = \frac{exp(V_i)}{\sum_{j \in J} exp(V_j)}$$ 3.3

where J is the full set of possible paths considered for the given traveler.

Dynamic Traffic Assignment

DynaMIT combines the models presented above to estimate and predict traffic network state and generate traffic information. For example, at the end of a time interval (e.g. 8:00 a.m.), DynaMIT performs state estimation using data collected during the last time interval (e.g. 7:55-8:00 a.m.). The estimated state of the network at 8:00 a.m. and the input toll rates are then used to generate state prediction for a future horizon (e.g. 8:00-8:20 a.m.).

State Estimation: Historical OD matrix, observed traffic surveillance data and information provided to travelers are passed to DynaMIT for state estimation. The estimated parameters (OD matrix and other key model parameters) would be obtained from the online calibration model, and then loaded into supply model. The estimated traffic network state and parameters would be used as inputs for state prediction.

State Prediction: Without measurement from the traffic surveillance system, the prediction modules uses the estimated network state as a starting point for prediction, and the most recently travel time guidance as trial strategy to solve the fixed-point problem. Autoregressive process is used to predict future OD matrix, and similar demand/supply models are used to obtain a revised travel time guidance. The module would run iterations until convergence in order to generate consistent information and guidance.

Figure 3.1 and Figure 3.2 illustrate the main modules of state estimation and prediction, and the impact of tolls rates for each module.

33

DynaMIT
State Estimation

Pre-trip Demand Simulation
Departure time and Route choice

Historical Database

Toll rate

Sensor data

Updated OD matrices

State estimation
OD and parameter calibration
En-route demand simulation
Supply simulation

Estimated traffic state
Calibrated parameters
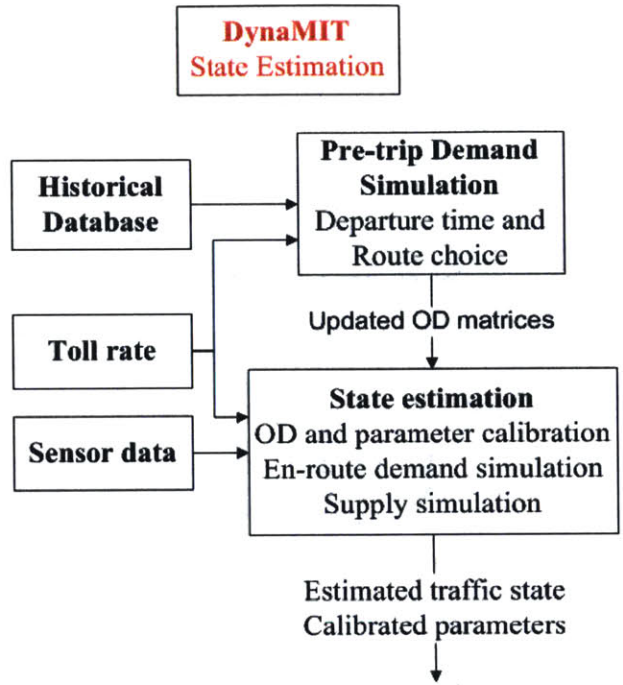
Figure 3.1 Road pricing in DynaMIT state estimation

DynaMIT
State Prediction

Demand prediction

Estimated traffic state
Calibrated parameters

Predicted OD

Information Generation

Pre-trip Demand Simulation

Toll Optimization

State Prediction
Demand simulation
Supply simulation

Predicted state

Return predicted state

Yes

Consistency?

No
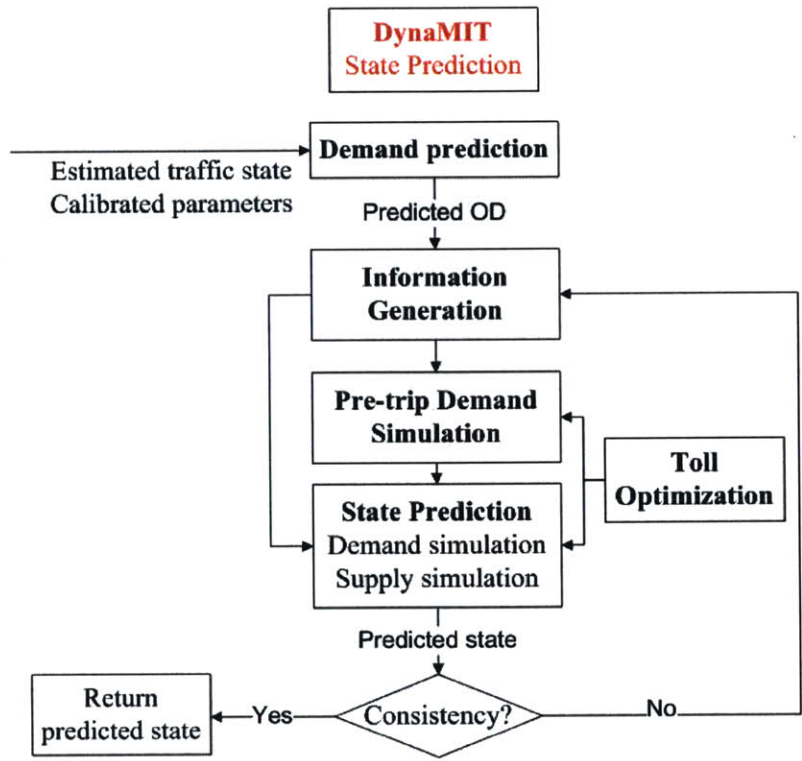
Figure 3.2 Road pricing in DynaMIT state prediction

3.2 Modelling approach

The proposed modeling approach is an *adaptive integrated simulation-optimization framework* for dynamic toll optimization that is depicted in Figure 3.3. Toll optimization interacts with state prediction in order to determine the optimal toll. The impact of toll on travel behavior is embedded through route choice models. The optimal toll value is sent back to the traffic surveillance system for the implementation of the updated toll.
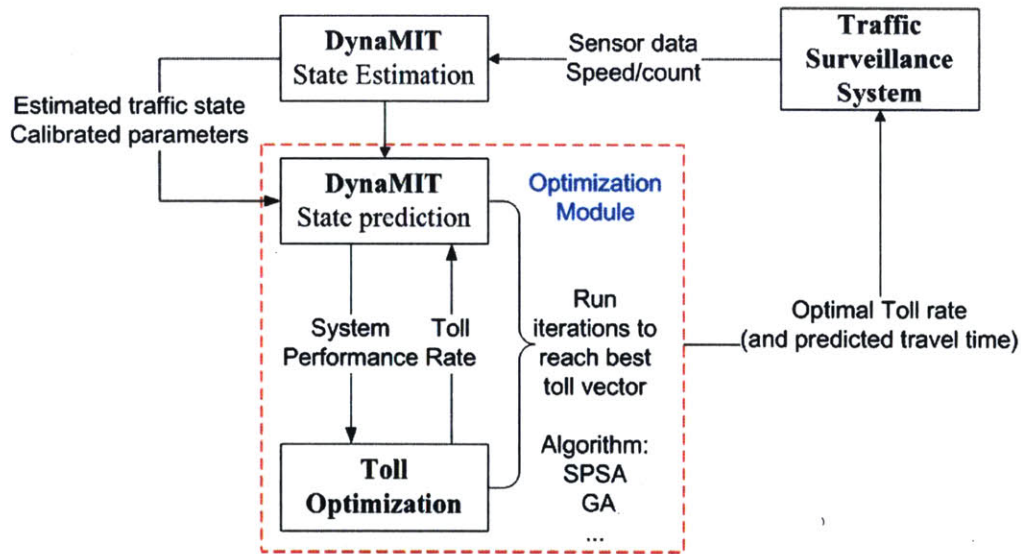
Figure 3.3 The adaptive integrated simulation-optimization approach

3.3 Optimization problem formulation

3.3.1 Rolling horizon scheme

The novelty of the proposed dynamic pricing methodology is that the optimization model is in a complete interaction with the prediction capabilities of the simulator as seen in Figure 3.3. It is developed in a rolling horizon manner so that as new information is received from the prediction module, the toll is re-optimized. We provide the terminology we use for the rolling horizon:

35

- **Tolling interval:** Defines how frequently the toll is optimized. It is flexible in our framework.

- **Optimization horizon:** Defines how far ahead we are optimizing. This is also flexible in the proposed framework.

- **Prediction horizon:** Defines how far ahead the prediction module of DynaMIT is performing traffic predictions. This is also flexible and there is a clear trade-off between the accuracy of the prediction and the length of the prediction horizon.

We provide the notation used for the toll optimization problem in Table 3.1. In Figure 3.4 we illustrate the rolling horizon toll optimization. The vector of decision variables for the toll is represented by $\beta = (\beta_1, ..., \beta_T)$. Since the optimization needs a certain computational time, the toll for the first tolling interval will not be implemented meaning that we will need to use the previously optimized toll value. The computational time for toll optimization is aimed to be less than one tolling interval. Therefore, in the figure the optimization horizon is shown to start with the next tolling interval and the toll for the first interval is represented by λ. We optimize the toll for the next T intervals and in the next tolling interval we will start another optimization. λ_2 will be equal to β_1^*, which is optimized in the first tolling interval.

We investigate two versions of the rolling horizon optimization:

- **Opt version 1 - Same toll during the optimization horizon:** In this version, we optimize a single toll value for a given managed lane segment during the optimization horizon. However, since we start another optimization routine in the next interval, implemented tolling at each interval may still be different. Referring to Figure 3.4, at the first tolling interval the optimized tolls will be equal to each other, i.e. $\beta_1^* = \beta_2^* = \cdots = \beta_T^*$. At the second tolling interval we will implement β_1^* and re-optimize the toll for the subsequent intervals in the optimization horizon.

- **Opt version 2 - Different toll for each tolling interval in the optimization horizon**: In this version, we optimize a different toll value for each tolling interval in the optimization horizon. This version is a generalization of the first version and the solution typically takes longer since the dimension of the search space increases.
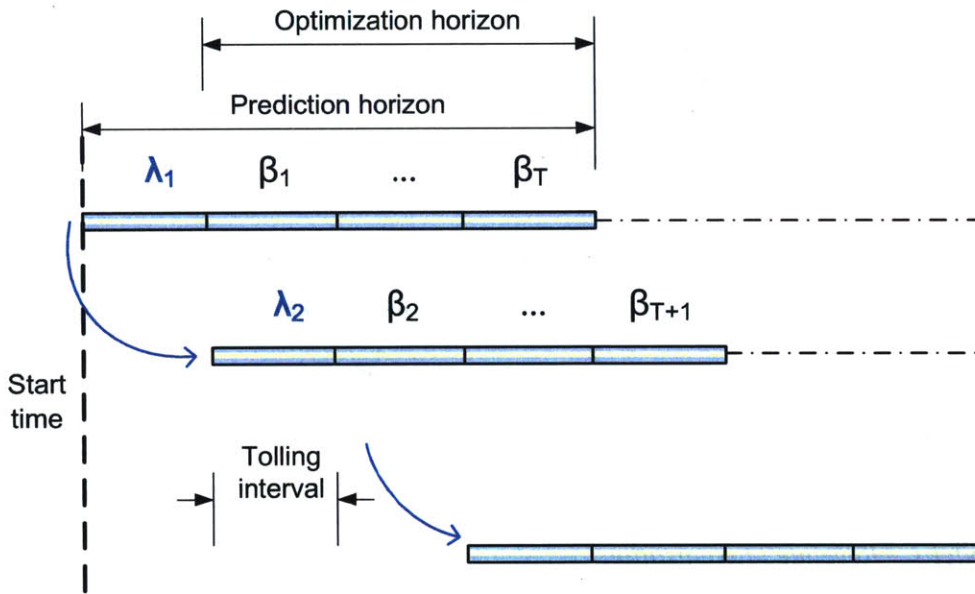
Figure 3.4 Illustration of the rolling horizon scheme.

3.3.2 Mathematical formulations

Table 3.1 Notation for toll optimization problem

Parameters	
I	Number of managed lanes (or tolling segments) indexed by i.
T	Number of tolling intervals in the optimization horizon indexed by t.
q_{cr}	Critical flow on the managed lane, which varies according to the total flow in order to maintain an efficient use of the managed lane.
v_{cr}	Critical speed on the managed lane in order to avoid congestion. This is a fixed value determined according to the network conditions.

α_v	Penalty on the deviation from the critical speed.
α_q	Penalty on the deviation from the critical volume.
α_β	Penalty on changing the toll with respect to the previous tolling interval.
Variables	
β_{it}	Decision variable - The toll rate on managed lane i at tolling interval t.
q_{it}	Number of vehicles entering managed lane i at tolling interval t as a function of the toll rate.
v_{it}	Average speed on managed lane i at tolling interval t as a function of the toll rate.

We present two formulations for the toll optimization problem. In the first formulation we maximize revenue and in the second one we consider maintaining certain traffic conditions on managed lanes while maximizing revenue. Note that the framework is flexible to accommodate different objective functions.

Model 1 – Revenue Maximization

We maximize total revenue without any constraints on the network conditions with the following model:

$$\max \sum_{i=1}^{I} \sum_{t=1}^{T} q_{it} \beta_{it} \qquad 3.4$$

$$s.t. \quad (v_{it}, q_{it}) = \text{DynaMIT}(\beta) \qquad 3.5$$

$$\text{LB} \leq \beta_{it} \leq \text{UB} \qquad 3.6$$

The objective function is the total revenue collected through the tolls on managed lanes. Constraints 3.5 maintain that the predicted speed and volume for each managed lane segment and tolling interval are provided by DynaMIT as a function of the toll vector. Since DynaMIT has various modules we cannot explicitly represent the function. Constraints 3.6 define the lower and upper bounds on the toll. Upper

bound is typically determined by the transport authority and the lower bound could be zero.

The presented formulation is the general case (*Opt version 2*) where each tolling interval in the optimization horizon has a specific toll variable. *Opt version 1* can be represented with a toll variable β_i so that each tolling segment i has a single toll value during the optimization horizon.

Model 2 – Revenue Maximization subject to Traffic Conditions

In this formulation, the revenue maximization is subject to conditions on the network. These constraints can be in various forms. Here we consider constraints on speed and volume on the managed lane. The model is given as follows:

$$\max \sum_{i=1}^{I} \sum_{t=1}^{T} q_{it} \beta_{it} \qquad \qquad 3.7$$

$$s.t. \quad v_{it} \geq v_{cr} \qquad \qquad 3.8$$

$$q_{it} \geq q_{cr} \qquad \qquad 3.9$$

$$(v_{it}, q_{it}) = \text{DynaMIT}(\beta) \qquad \qquad 3.10$$

$$\text{LB} \leq \beta_{it} \leq \text{UB} \qquad \qquad 3.11$$

The objective function 3.7 and constraints 3.10-3.11 are the same as before. Constraints 3.8 and 3.9 maintain conditions on speed and volume respectively. The speed needs to be above a critical value on the managed lane so that the congestion is avoided. Volume is also aimed at being above a critical value in order to avoid very little usage of the managed lane. The two constraints together aim at keeping the managed lane uncongested, i.e. running above the desired speed, but at the same time efficiently utilized, as illustrated in Figure 3.5.

Here it should be noted that v_{cr} should vary over time depending on the total demand level, and should be calibrated in real time. The main purpose of introducing

such a constraint is to evaluate the system's reaction to different restrictions on operation performance, and to test if the system is running the optimization in an expected manner. When applying on a real-world network, the specification of constraints will be further studied to fit the need from road operators.
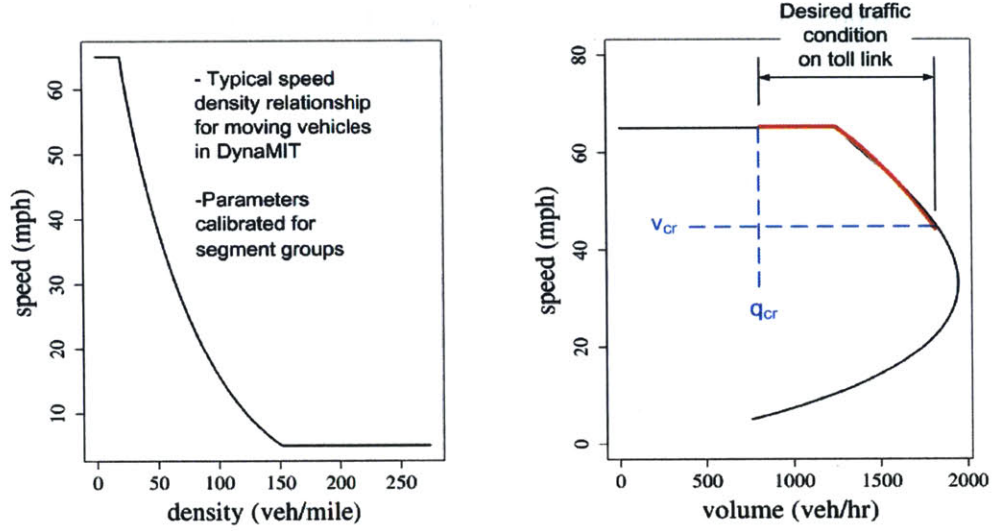
Figure 3.5 Illustration of traffic condition constraint

The model presented in 3.7-3.11 introduces hard constraints on the deviation from critical volume and speed. However, it would be more realistic to allow certain deviations with penalties since it is not easy to satisfy the constraints especially in peak hours. Constraints 3.8 and 3.9 can be moved to the objective function with weights analogous to lagrangian multipliers such that when the constraint is not binding there is no penalty. We present the modified model as follows:

$$\max \sum_{i=1}^{I} \left[\sum_{t=1}^{T} q_{it} \beta_{it} + \alpha_v \frac{1}{T} \sum_{t=1}^{T} \min(v_{it} - v_{cr}, 0) + \alpha_q \frac{1}{T} \sum_{t=1}^{T} \min(q_{it} - q_{cr}, 0) \right] \quad 3.12$$

$$s.t. \quad (v_{it}, q_{it}) = DynaMIT(\beta) \quad 3.13$$

$$LB \leq \beta_{it} \leq UB \quad 3.14$$

The objective function 3.12 consists of three terms. The first term is the total revenue obtained in the entire optimization horizon as before. The second term is the

penalty on the speed deviation meaning that if the speed on a managed lane goes below the critical value it will be negative and reduce the objective function value. Similarly, the last term is the penalty on the volume. The critical volume varies based on the total flow such that it will increase as total flow increases in order to keep a balanced level of traffic volume on the managed lane and the free lane. The presented *Model 2* formulation adopts *Opt version 2*, and *Opt version 1* can be represented with a toll variable β_i.

3.4 Testing approach

3.4.1   Close-loop testing framework

In the optimization framework, DynaMIT is used as a tool to generate the optimal tolling strategy. In order to evaluate the tolling strategy, there are generally two approaches: field test and simulation-based evaluation. While field tests are expensive and technically difficult to implement, the simulated output are usually analyzed to evaluate the control strategy. One potential deficiency of simulated-based evaluation lies in the fact that the strategy is both generated and evaluated in the same simulator. The bias results from the same traffic flow dynamic and traveler behavior models used to generate and evaluate strategies, and this setting would invariably overestimate the traffic system improvement from the strategy.

In order to test our modeling approach, we use a microscopic traffic simulator, MITSIM, which represents the real-life traffic surveillance system given in Figure 3.6. Therefore, a *closed-loop* approach is built in a loop with a microscopic simulator and a mesoscopic simulator. The software interface that developed to operationally link DynaMIT and MITSIM provides a seamless facility for the exchange of information between DynaMIT and MITSIM in real time (Ben-Akiva et al., 2002a; Rathi et al., 2008).
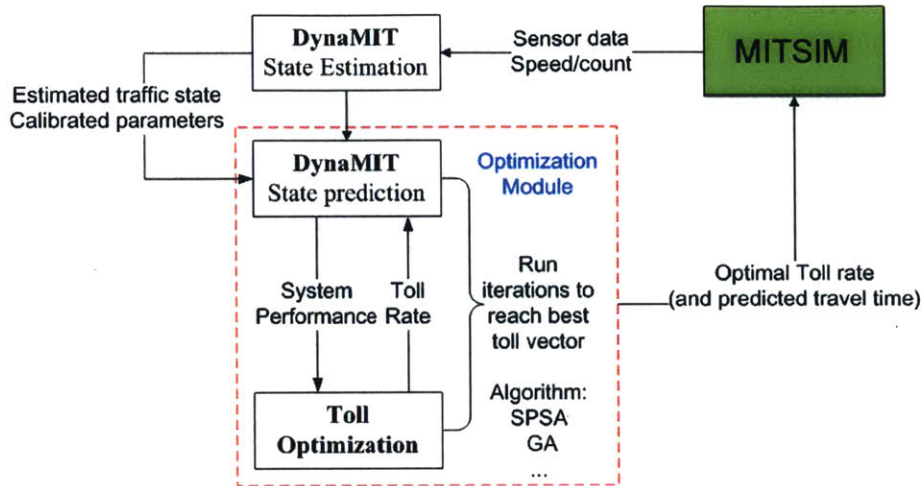
Figure 3.6 Close-loop testing framework

Interaction between components

(1) MITISM runs the simulation for the current interval, and send to DynaMIT traffic surveillance data (e.g., loop detector count and speed), accident report, and traffic messages disseminated to the drivers.

(2) DynaMIT receives the latest traffic data and uses it for state estimation and state prediction.

(3) The optimization module optimizes the toll in interaction with the state prediction based on pre-defined objective functions and constraints.

(4) The travel guidance information and the generated optimal toll rate are then sent back to MITSIM.

(5) MITSIM runs traffic simulation for next interval with the strategy and travel guidance from DynaMIT, and again outputs new traffic surveillance data to be fed back to DynaMIT.

Experiment set-up procedure

(1) Generate network file in both MITSIM and DynaMIT, including network topology, specification of nodes, links, segments, lane, lane connections, etc.

(2) Calibrate MITSIM against real-world filed data.

(3) Run MITSIM with calibrated parameters and output simulated traffic surveillance data.

(4) Calibrated DynaMIT against the MITSIM simulated output.

The full calibration process for a real-world traffic network will be described in the case study section.

3.4.2   MITSIM

MITSIM is a microscopic traffic simulation model that evaluates the impacts of alternative traffic management system designs, traveler information systems, public transport operations, and various ITS strategies at the operational level and assists in their subsequent refinement. MITSIM can evaluate systems such as advanced traffic management (ATMS) and route guidance. MITSIM was developed by MIT's Intelligent Transportation Systems (ITS) Program (Q. Yang et al., 2000; Q. Yang and Koutsopoulos, 1996).

MITSIM represents networks at the lane level and simulates movements of individual vehicles using car-following, lane changing, and traffic signal response logic. Probabilistic route choice models are used to capture drivers' route choice decisions in the presence of real time traffic information provided by route guidance systems. MITSIM is designed as a testbed for evaluating traffic management systems.

Driving behavior model

The driving behavior models deal with tactical and operational driving decisions, mainly acceleration and lane changing. Driver behaviors are mainly controlled by two models: acceleration model and the lane-changing model (Ben-Akiva et al., 2002b).

The acceleration model determines the longitudinal movement of a vehicle under three possible regimes: free-flowing, car following and emergency, while the lane-changing model captures a lane-changing action with three levels of decision-making: the decision to change lane, the choice of lane to change to and the gap acceptance

behavior to perform the lane change.

Travel behavior model

The travel behavior models include both pre-trip and en-route path choices. Drivers in MITSIM may have either predefined paths or compute them online. Depending on whether set of paths is predefined for the traveler or not, a path-based route choice model or a link-based route generation model would be used (Ben-Akiva et al., 2010b).

The effects of traveler guidance and tolls on route choice are captured by the route choice models. MITSIM classified the drivers to be informed and uninformed, where the informed drivers would make route choices based on updated travel time which incorporates real time traffic condition. If en-route traveler information is available to drivers, they will update their route choice whenever the new information is sent to them. There is also a diversion dummy variable penalizing drivers' switching from the previous route. Uninformed drivers rely on habitual or historical travel time to make route choice decision.

3.5 Summary

This chapter first described how to take into account the impact of road pricing on drivers' route choice in traffic simulators. Then the framework is developed for optimization of real time tolling strategy. The mathematical formulation of optimization problem is proposed with multiple forms of objective functions. A testing approach is proposed to incorporate DynaMIT and MITSIM running in a close-loop framework.

4 Application on synthetic network

4.1 Overview of synthetic network

The closed-loop approach presented earlier will be the platform for generating simulation experiments for different case studies. In this chapter, the aim is to have an understanding of the impact of toll optimization on network conditions through different scenarios. Therefore, we present results only with DynaMIT. We use a simple solution algorithm for the numerical experiments in this chapter such that we have an exhaustive search with $0.05 of increments.

For all the formulations considered in this chapter, we introduce a small penalty on changing the toll with respect to the previous tolling interval if the revenue gain is below a predetermined threshold value. This term, $\alpha_\beta \left| \beta_{it} - \beta_{i(t-1)} \right|$, is introduced in order to avoid fluctuations on the toll due to very small changes in the objective function value. The penalty value, α_β, is arranged such that a change in toll is avoided when the revenue gain is below the threshold. We assume a value of 70, which means that in order to have a change of $0.05 in the toll, it should bring at least $3.5 of revenue during the prediction horizon.
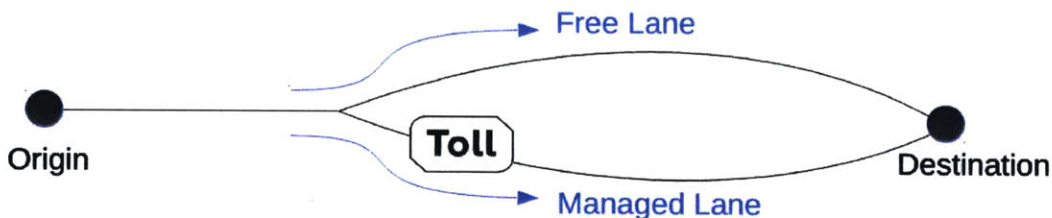

Figure 4.1 Network representation for the simple example.

The network chosen for the numerical experiments is presented in Figure 4.1. There are two possible paths from origin to destination. One path goes through a Free Lane (FL) and the second one goes through a Managed Lane (ML), where users need to pay a toll in the entrance after the bifurcation. There are two lanes on both FL and

45

ML and the capacity is assumed to be 3900 veh/hr for each. Maximum allowed speed on both the FL and the ML is 70 mph. The simulation horizon is 6:30-9:00, first 30 min being the warm-up period. The toll is optimized every 5 min, meaning that the tolling interval is 5 min. The prediction horizon is considered as 20 min. The optimization horizon is considered as 15 min so that the toll is optimized for 3 tolling intervals starting from the next tolling interval as illustrated in Figure 3.4.

The demand profile is assumed to be increasing such that for the four consecutive 30 min intervals during 7:00-9:00, it is 3600, 3900, 4200 and 4500 veh/hr respectively as in Table 4.1. Such a profile is considered in order to see the change in toll values as a reaction to increasing demand. As mentioned before, VOT is assumed to be distributed across the population and the mean value is assumed to be 15 $/hr.

Table 4.1 Input demand profile

Time interval	Demand (veh/hr)
6:30-7:00 (warm-up)	3600
7:00-7:30	3600
7:30-8:00	3900
8:00-8:30	4200
8:30-9:00	4500

4.2 Results

We present different experimental results in the remainder of this section. First of all we compare the results obtained with *Model 1* and *Model 2*. Secondly, we analyze the impact of different VOT distributions, namely different mean values. Finally, we compare *Opt version 1* and *Opt version 2* in order to see if there is a significant loss when optimizing a single toll value for the entire optimization horizon for each tolling segment. Note that except this last experiment all the results are obtained with *Opt version 1*.

Model 1 vs Model 2

We first present results with *Model 1* which has an objective function of revenue maximization. In Figure 4.2 and Figure 4.3 we present the DynaMIT simulated flow and speed, respectively. As expected, the flow on FL is increasing and the speed is decreasing as a consequence of the increasing demand profile. On the other hand, ML is kept running at the maximum speed with a relatively steady flow. This is achieved by increasing the toll as illustrated in Figure 4.4. In the first hour, the toll is steady at $0.5 and increases gradually to $0.65 and $0.7. Note that, the penalty on changing the toll avoids fluctuations.
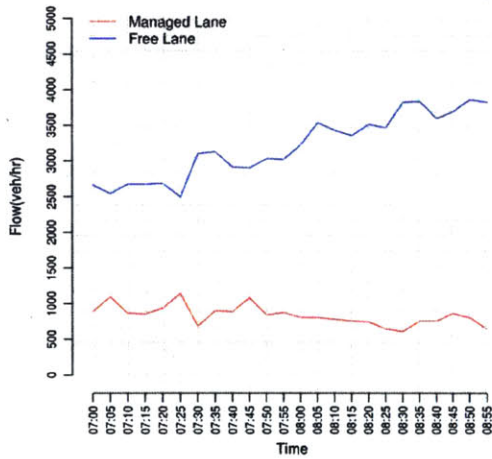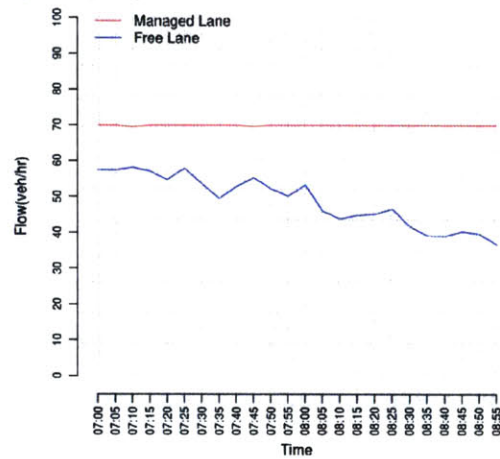

Figure 4.2 Simulated flow (Model 1)

Figure 4.3 Simulated speed (Model 1)
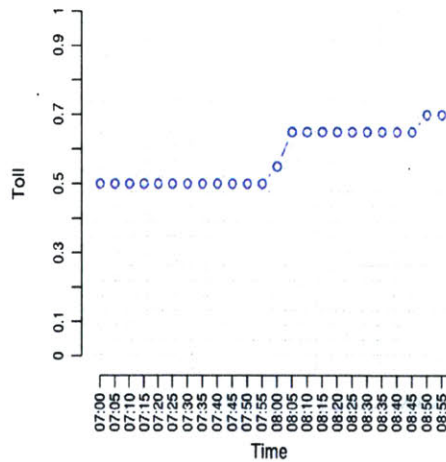
Figure 4.4 Optimized toll values (Model 1)

Next, we analyze the results of *Model 2,* which considers traffic conditions while maximizing revenue. The critical speed on ML is assumed to be 65 mph and the critical volume on ML is assumed to be 1200 veh/hr between 7:00-8:00 and 1500 veh/hr between 8:00-9:00. As previously mentioned, the critical volume is adjusted based on the total flow in order to better adapt to changing demand patterns.

The simulated flow and speed are presented in Figure 4.5 and Figure 4.6, respectively. It is observed that the condition on FL is better compared to *Model 1* mainly because the volume on ML is controlled to be above the critical value. ML is still around the maximum speed but goes below that few times during the simulation horizon. The optimized tolls are lower compared to *Model 1* as presented in Figure 4.7. The average toll value is around $0.4. Therefore the resulting total revenue is lower compared to *Model 1* as presented in Table 4.2 and Table 4.3 under *Setting 1*. Figure 4.8 displays the objective function for different toll values for the tolling interval 08:25-08:30. If we were to maximize revenue only, the optimal toll for this interval would be $0.65. However, in the existence of the speed and volume penalties, the optimal value shifts to $0.4 as shown in the figure. The speed penalty decreases and the volume penalty increases as toll increases.
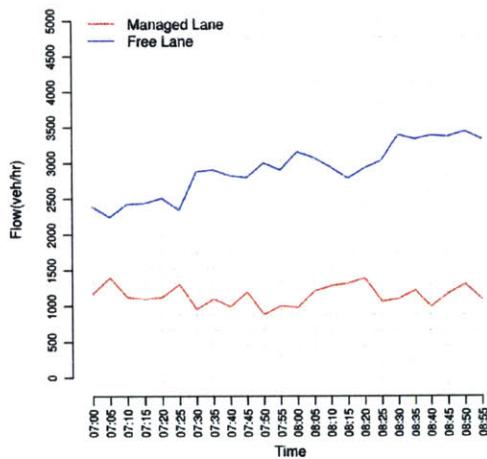
Figure 4.5 Simulated flow (Model 2)
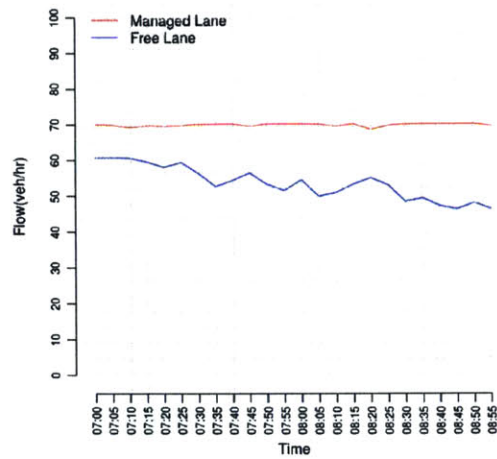
Figure 4.6 Simulated speed (Model 2)
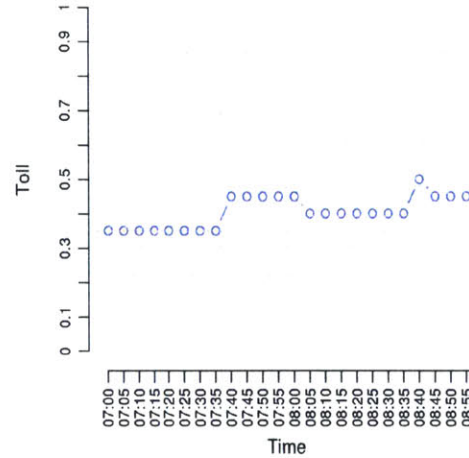
Figure 4.7 Optimized toll values (Model 2)
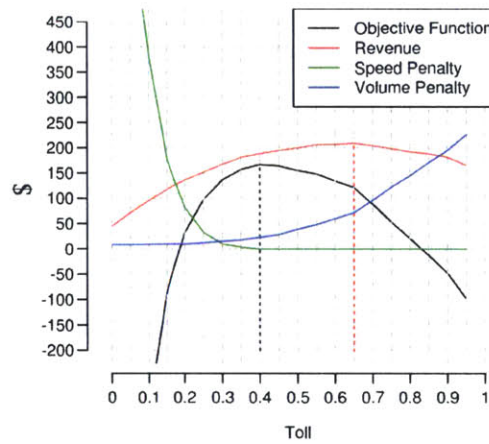
Figure 4.8 System performance at different toll rates.

Having constraints on speed and volume, which are introduced as penalties in the objective function, results with lower revenue as expected. Since the aim of the penalties is to have an uncongested and a better utilized ML, the throughput on ML increases without leading to significant speed decreases, which in turn alleviates the conditions on FL.

The results of *Model 2* depend on the values of penalties and critical volume and speed. When the critical volume on ML is increased or the penalty on volume is increased, optimized tolls go lower and as a result the revenue is lower. On the other hand, as the critical speed on ML is increased or the penalty on speed is increased, the tolls need to be higher in order to maintain speed conditions and a higher revenue is

obtained. We present results with additional penalty settings in Table 4.3 for *Model 2* for illustration purposes. When we compare *Setting 1* and *2*, we see that an increased penalty on volume results with a lower revenue since more drivers need to be attracted to ML. When we compare *Setting 2* and *3*, the increase in speed penalty results with higher revenue since the tolls are optimized at higher values in order to have fewer drivers on ML.

Table 4.2 Revenue and Throughput for Model 1

	Model 1	
	Opt version 1	*Opt version 2*
Average toll ($)	0.580	0.572
Total revenue ($)	946	947
ML throughput (veh)	2161	2163
FL throughput (veh)	7716	7716

Table 4.3 Revenue and Throughput for Model 2

	Model 2 *(Opt version 1)*				
	Setting 1 $\alpha_v = 3.0$ $\alpha_q = 1.5$ VOT=15$/hr	**Setting 2** $\alpha_v = 3.0$ $\alpha_q = 3.0$ **VOT=15$/hr**	**Setting 3** $\alpha_v = 6.0$ $\alpha_q = 3.0$ VOT=15$/hr	**Setting 4** $\alpha_v = 3.0$ $\alpha_q = 1.5$ **VOT=10$/hr**	**Setting 5** $\alpha_v = 3.0$ $\alpha_q = 1.5$ **VOT=20$/hr**
Average toll ($)	0.404	0.344	0.350	0.258	0.540
Total revenue ($)	915	864	869	605	1219
ML throughput (veh)	2840	3140	3124	2842	2821
FL throughput (veh)	7039	6740	6755	7037	7058

Impact of VOT

In this section we analyze the impact of lower (10$/hr) and higher (20$/hr) VOT with *Model 2*. In Table 4.3, *Setting 1* corresponds to the regular case with 15$/hr. *Setting 4* and *5* represent the lower and higher VOT cases, respectively. When VOT is lower the optimized toll values are lower since ML is less attractive to drivers. When the VOT is higher the tolls increase which results with higher revenue. Across the three versions, the throughput of ML is similar which is intuitive as the optimization model decides on the tolls based on VOT for matching a similar target of attracted drivers.

50

These results show that different behavioral parameters are reflected in the results through the embedded choice models in DynaMIT.

Same vs Different Toll during the Optimization Horizon

Finally, we analyze the results with *Opt version 1* and *Opt version 2* using *Model 1*. The aim of the analysis is to see how much we lose when we assume a single toll for the entire optimization horizon.

It turns out that for our simple example the two versions have quite similar network conditions and total revenue. *Opt version 1* has a total revenue of $945.5 with an average toll of $0.580 and *Opt version 2* generates a revenue of $947.4 with an average toll of $0.572, as illustrated in Table 4.2. When we think about the needed computational time, if we have 10 possible toll values for each tolling interval in our exhaustive search, *Opt version 1* has a dimension of 10. On the other hand, this is $10 \times 10 \times 10$ for *Opt version 2* since the toll is different for each tolling interval in the optimization horizon (assuming 3 tolling intervals). Since *Opt version 2* will always need more computational power and the results are not significantly different, *Opt version 1* might be preferred for real time studies.

4.3 Summary

The results suggest that the framework generates consistent results. An increasing demand profile is considered and this results in an increasing pattern of optimized tolls as expected. When constraints are introduced to maintain traffic conditions, the resulting revenue is lower compared to pure revenue maximization. However, such constraints alleviate network conditions in general. Furthermore, when drivers have a higher willingness to pay, the optimized tolls are higher as expected.

5 Case study on NTE network

In this chapter, a case study on NTE TEXpress Lanes network in Texas is presented to demonstrate the application of the optimization framework on a real-life highway corridor network.

5.1 Overview of NTE network

5.1.1   Network

The North Tarrant Express project developed a 13.3-mile TEXpress Lanes. Drivers can travel the NTE TEXpress Lanes from I-35W across I-820 (Northeast Loop) to SH-121/183 (Airport Freeway), just beyond the SH-121 split, with no stop-and-go traffic along the way. The entire road network for NTE TEXpress network is shown in Figure 5.1. The corridor network is extracted and modeled in both DynaMIT and MITSIM. The simulated network includes general purpose lane, managed lane, on-ramp/off ramps. The network representation is manually generated from the original computer-aided engineering drawings. The center lines of each road segment are used to determine the starting point, ending point and curvature of each segment.

Figure 5.1 NTE TEXpress Lanes network map

There are totally 146 nodes and 167 links in the highway network. Each link is divided into several segments for an accurate representation of network geometry, and

there are totally 266 segments. Based on the network topology, 246 OD pairs are selected to make sure that each origin and destination combination is feasible on the network.

5.1.2   Tolling in practice

The highway operator is currently implementing real time tolling based on professional intuition so as to keep the traffic flow on managed lane no lower than 50 mph. As traffic levels and demand increase, the toll price changes to keep vehicles moving. Once traffic volumes drop, the price goes down[11]. So generally the toll is higher during rush-hour periods on the weekdays, and lower during non-peak periods at other times of the day and on the weekends. Table 5.1 shows an example of average toll rates for Eastbound I-820 East Entrance Ramp during different intervals in the morning peak hours. The general pattern is a higher toll value during 6:30-8:30 AM on weekdays, and lower tolls during off-peak periods.

Table 5.1 Average tolling rate for Eastbound I-820 East Entrance Ramp (G3)

	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
06:00 AM	$0.95	$3.00	$3.08	$3.08	$2.94	$2.93	$0.91
06:30 AM	$0.95	$3.68	$3.70	$3.70	$3.51	$3.55	$0.91
07:00 AM	$0.95	$3.99	$4.06	$4.04	$3.75	$3.83	$0.91
07:30 AM	$0.95	$4.05	$4.17	$4.11	$3.82	$3.88	$0.91
08:00 AM	$0.95	$4.00	$4.10	$4.08	$3.82	$3.85	$0.91
08:30 AM	$0.95	$3.89	$4.02	$4.04	$3.83	$3.82	$0.91
09:00 AM	$0.95	$2.94	$3.06	$3.18	$2.96	$2.90	$0.95
10:00 AM	$0.95	$1.89	$1.94	$1.97	$1.88	$1.83	$0.95

The NTE TEXpress Lanes are comprised of two tolling segments:

- Toll Segment 1 (indicated in purple in Figure 5.1) runs from 35W to the Northeast Loop / Airport Freeway interchange, near North East Mall
- Toll Segment 2 (indicated in blue in Figure 5.1) begins at the Northeast Loop / Airport Freeway interchange and extends to Industrial Boulevard

[11] http://www.ntetexpress.com/pricing/how-pricing-is-determined

53

The tolling strategy is segment-based, and implemented in the following ways,

- The toll rates are presented on a pricing sign before the toll gantries so that travelers have enough time to decide whether to enter the managed lane or not Figure 5.2 shows an example of entrance ramp.

- Pricing sign will also be presented on managed lane before entering the next toll segment, so that travelers may choose to exit managed lane if they are not willing to pay the toll rate for next segment. Once travelers enter a tolling segment, they are charged for that segment.

- Full rate is charged at toll gantries located at the beginning of each toll segment, and a predefined discount factor is used to calculate the toll rate at other gantries along the toll segment. See Table 5.2 for discount factors at each toll location. This factor is usually fixed within a day, and cannot be changed real time as toll rates.

Figure 5.2 Illustration of Airport Fwy North Entrance, Exit 23A[12]

---

[12] http://www.ntetexpress.com/maps/nte-texpress-lanes-entrance-ramp-maps

Table 5.2 Discount Factors by Gantry

Discount Factors by Gantry							
Segment 1 Westbound		Segment 1 Eastbound		Segment 2 Westbound		Segment 2 Eastbound	
gantry	factor	gantry	factor	gantry	factor	gantry	factor
5	0.25	3	1.00	16	1.00	9	1.00
8	0.75	6	0.50			11	1.00
10	1.00	7	0.50			12	0.75
						13	0.50

5.1.3   Data

The highway road operator CINTRA provided three types of data for the week of June 7th -June 13th, 2015 as follows.

Sensor data

Flow and speed data are obtained from the MXD detectors located along general purpose lane, managed lane and ramps in the network. The distance between the detectors is approximately 2000 feet. The data is reported to the control center every minute, and there are totally 135 sensors available. Figure 5.3 gives an example of such data provided by a specific sensor.

Manual inspection is conducted to identify the inconsistent and erroneous data before applying the off-line calibration. Firstly, those sensors with many "NA" or zero values for speed and flow are eliminated. Next, some malfunctioning sensors are identified and eliminated from further use. Specifically, if there is an on-ramp/off-ramp between sensor i and sensor j, and another sensor k located at the ramp, then the law of flow conservation can be used to check if the three sensors demonstrate inconsistency in counts, and further investigation would identify which one of them is erroneous.
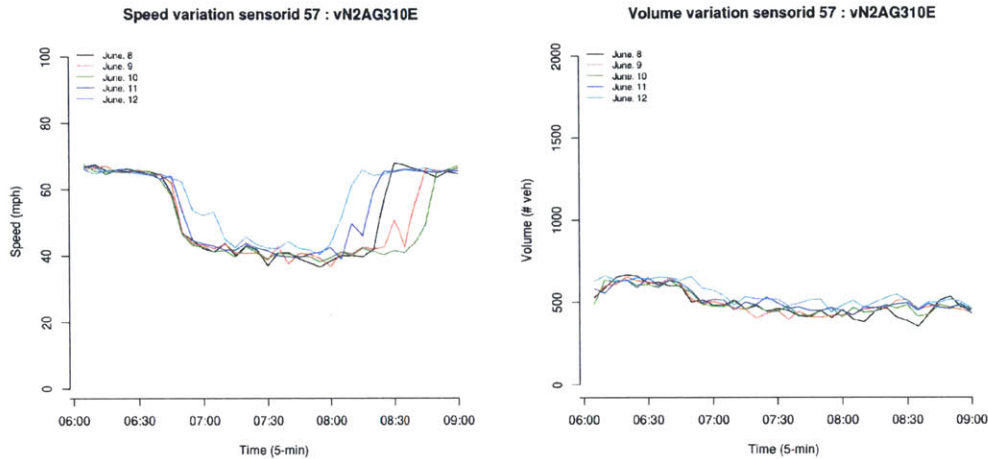
Figure 5.3 Five-minute speed and volume for sensor no.57

The aggregated sensor data can also provide insights on the demand pattern of the overall network and the variation of demand levels across weekdays. Figure 5.4 illustrates the total traffic volume change aggregated by each segment type. Figure 5.5 illustrates the average sensor speed change aggregated by each segment type. For example, E1G represents eastbound segment 1 general purpose lane, while W2M represents westbound segment 2 managed lane.

Generally, the morning peak hours last from 6 AM to 9 AM, while the evening peak hours start early afternoon, and end around 8:30 PM. The peak direction is eastbound in the morning and westbound in the evening. The operation speed is stable and higher than 60 mph for the managed lane. On the contrary, there is significant speed reduction observed on general purpose lane, with the average speed decreasing below 40 mph during congestion.
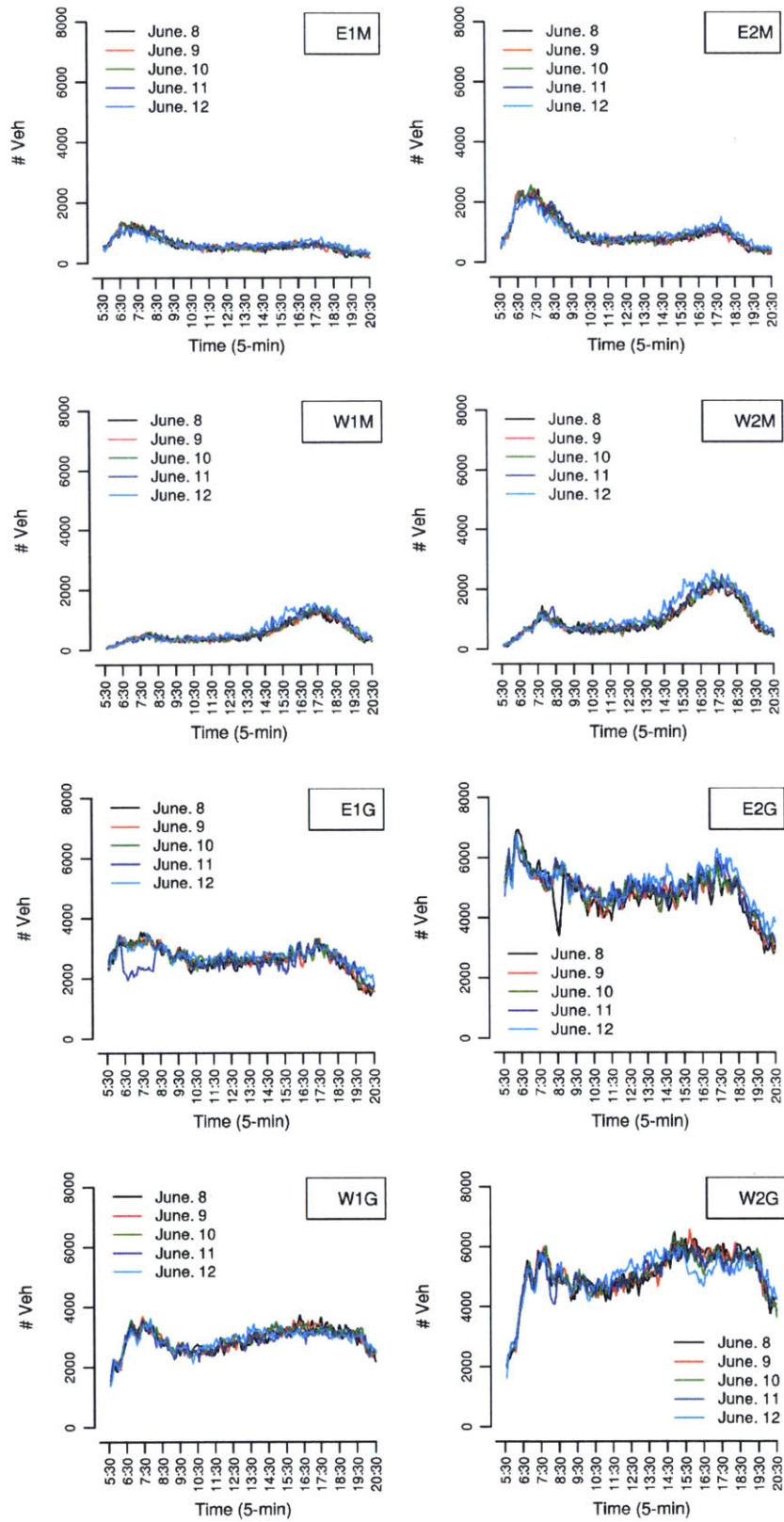
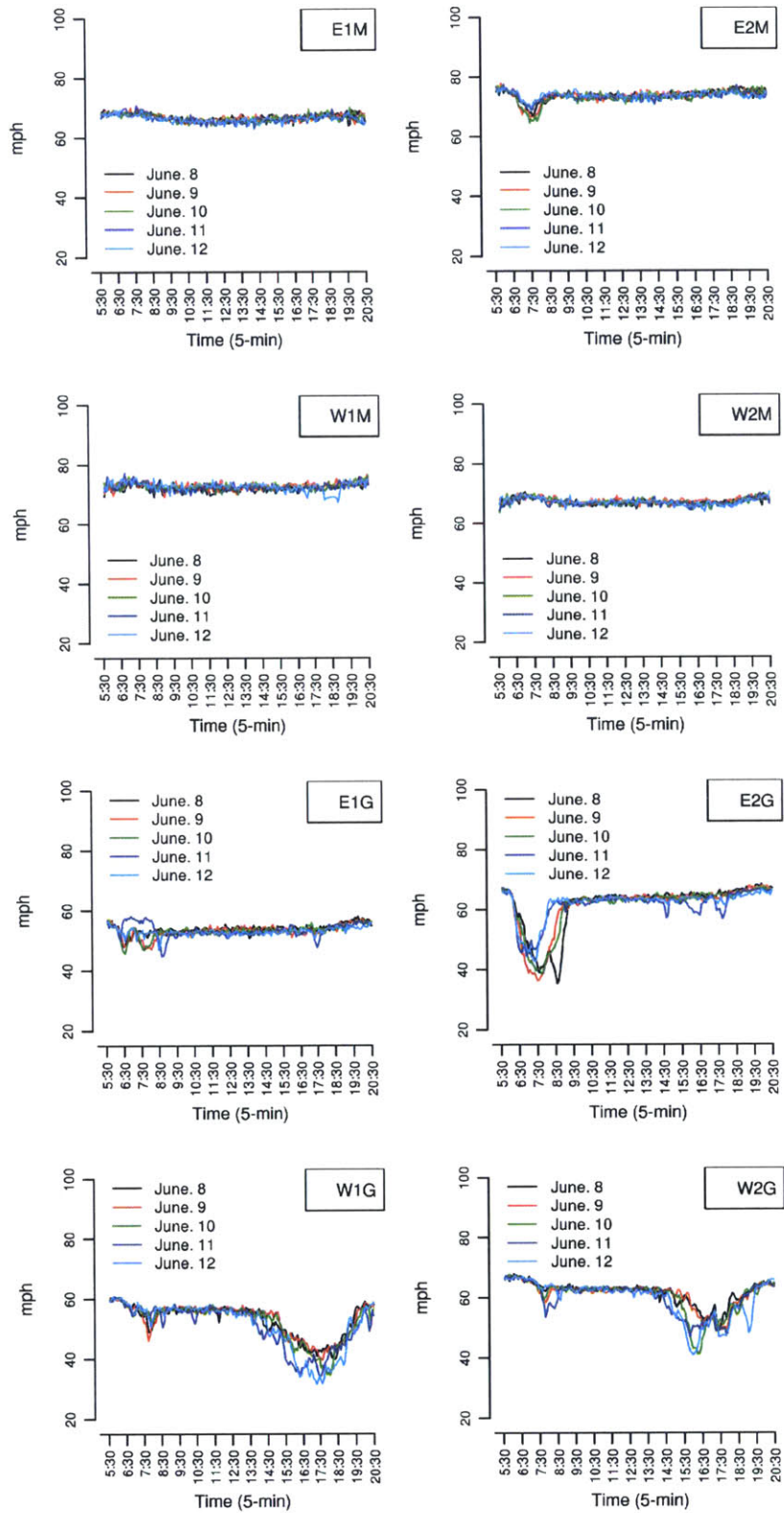Figure 5.4 Variation of total sensor count for each segment group

57

Figure 5.5 Variation of average sensor speed for each segment group

58

AVI data

AVI data is provided by the road operator and captures the information for vehicles equipped with transponders. Most of the AVI tag detectors are located at the on-ramps and off-ramps connecting general purpose lane, and a few of them connecting managed lane. See Figure 5.6 for locations AVI tag detectors across the network. Each record of the dataset includes vehicle transponder ID and time stamp. Although AVI data is collected whenever vehicles equipped with transponders pass a tag detector, it is not available to the control center in real time. Instead, AVI data for the entire day is communicated back to control center at the end of each day.
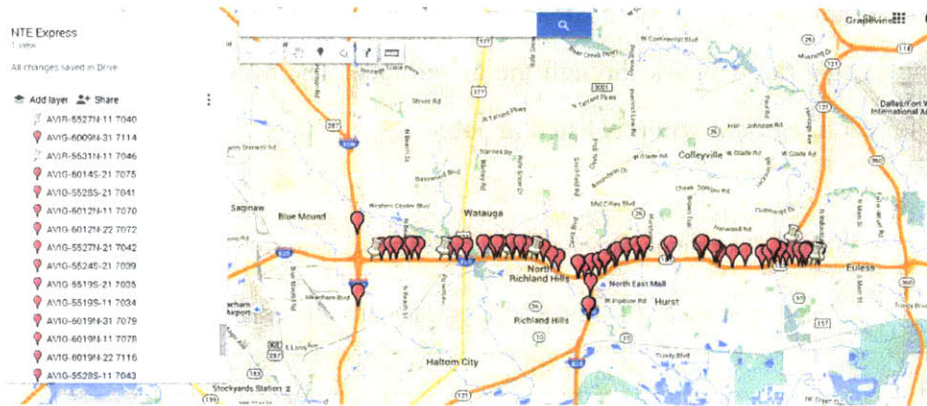
Figure 5.6 Distribution of AVI tag detectors

Before utilizing AVI data, it is helpful to investigate the market penetration rate of transponders for traveler using the network. The segments with both sensors and AVI tag detectors are selected to calculate penetrate rate for morning period, and the results are average by the toll segment where the AVI tags detectors are located. The results in Figure 5.7 shows the variation of penetration rate on June 8, and generally the rate on managed lane is higher than general purpose lane and ramps.
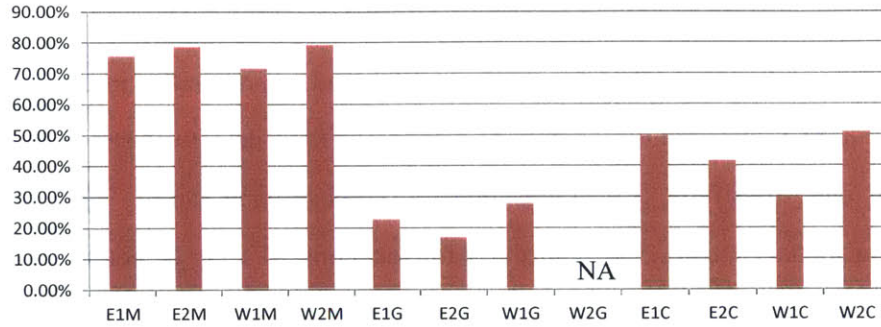
Figure 5.7 Morning peak penetrate rates grouped by segment type

To further make use of the AVI data in OD expansion, the penetration rates are further aggregated to the node level. The definition of penetration rate for an origin/destination node is defined as the proportion of vehicle with transponders entering/exiting the network through the given node. The average penetration rate of nodes for each segment group is listed in Table 5.3.

Table 5.3 Penetration rates at nodes grouped by segment group and weekday

| Segment Group	June. 8th	June. 9th	June. 10th	June. 11th	June. 12th
E1C	30.29%	30.51%	30.57%	28.89%	28.57%
E2C	37.55%	38.50%	41.48%	38.37%	36.29%
W1C	35.37%	35.98%	35.53%	35.37%	33.08%
W2C	38.60%	39.50%	39.14%	40.43%	36.57%

Transaction data

Transaction data is collected at each of the 11 toll gantries. Each record of the dataset includes vehicle plate ID, vehicle transponder ID (if equipped), Gantry ID, vehicle type, time stamp and toll rate. It should be noted that the proportion of vehicles captured by toll gantries is approximately 99%, and both equipped and non-equipped vehicles are captured.

Since transaction data captures each individual vehicle, they can be treated as both sensor data (i.e. provide 1-min count data) and AVI data (i.e. provide vehicle tag ID and time stamp), so that the three types of data are combined as complementary

60

datasets. Figure 5.8 and Figure 5.9 illustrate traffic flow detected at every toll gantry during the peak hours on June 8, while Figure 5.10 shows the real time toll rates for each toll segment. Note that in real-life implementation, there is discount rate, overcharge or exemption for particular types of vehicles. For example, heavy trucks are subject to higher charges compared to single occupancy vehicles. High occupancy vehicles can pay less if they activate HOV status for the desired HOV periods via web site or the mobile app before entering the TEXpress Lanes.
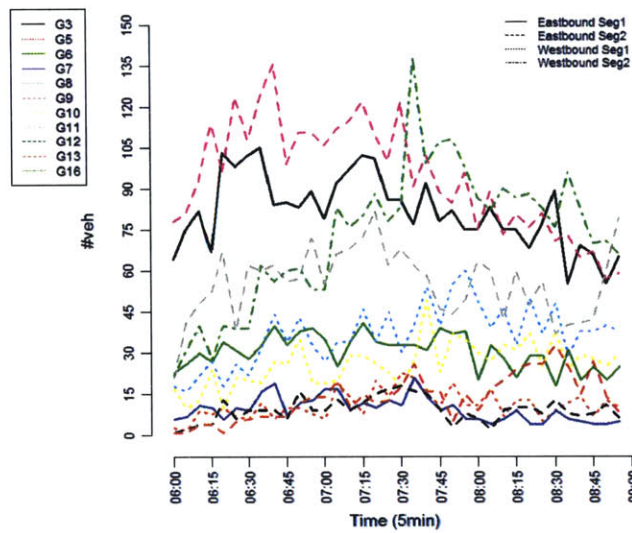

Figure 5.8 Volume detected at toll gantries during 6:00-9:00 on June 8th
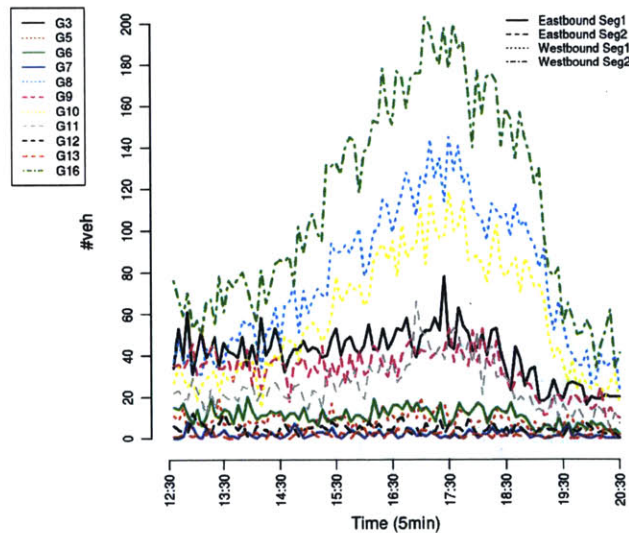

Figure 5.9 Volume detected at toll gantries during 12:30-20:30 on June 8th
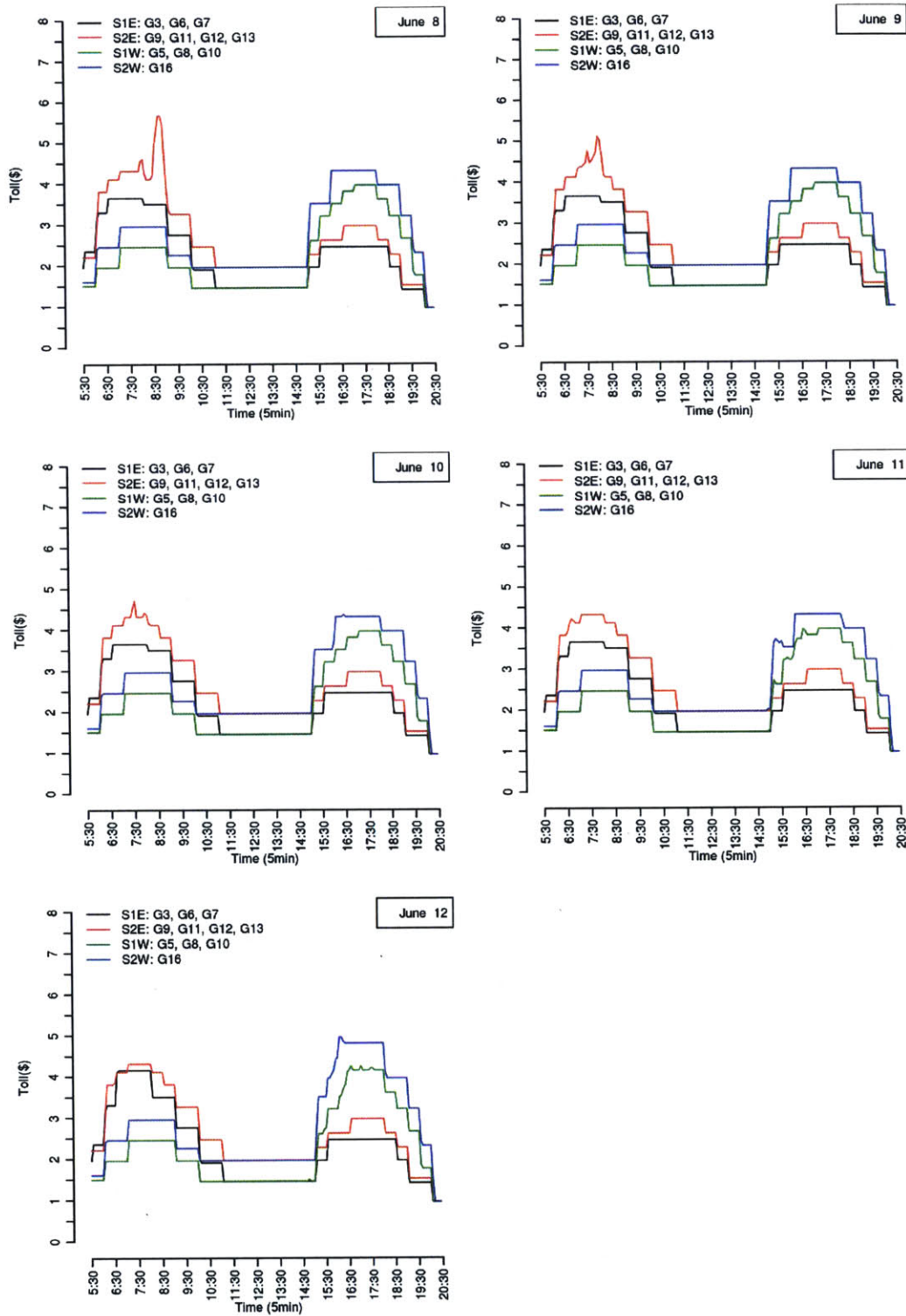
61

Figure 5.10 Five-minute toll rates for each toll segment

5.2 Offline calibration methodology

5.2.1 Overview of model calibration for DTA model

In order to utilize DTA models to evaluate the impact of traffic control strategies on a real-life network, the first fundamental step is to replicate the real-world demand patterns and traffic dynamics in an accurate way inside the traffic simulator. This requires a systematic approach to estimate all the key parameters in the DTA model based on observed traffic data, and this procedure is called calibration.

Based on the frequency and time of calibration, there are two main types of model calibration: offline and online. Offline calibration uses historical traffic database (e.g., planning OD matrix, traffic sensor data, etc) to estimate the parameter values so that the simulator can replicate the average traffic conditions of the network. However, some long-term effects, like type of day and seasonal effects, may compromise the estimation results. Online calibration differs from offline in that it uses the offline calibration results as historical database, and adjusts the parameter values based on real time traffic surveillance data, so that the short-term factors like weather and incidents are better captured.

The parameters from both the demand and supply sides need to be calibrated against the observed data. The typical parameters on demand side include time-dependent OD matrix, travel behavior model parameters (e.g., route choice), etc. On supply side, modelers are interested in calibrating driving behavior model parameters (e.g., car-following model, lane changing), segment capacity, speed-density function model, etc.

To capture the interaction between demand and supply components, latest research uses simultaneous demand-supply calibration framework to formulate the offline calibration problem (Balakrishna et al., 2007). The objective function is the weighted sum of various types of deviations between observed traffic measurement and simulated traffic measure (e.g., sensor count, speed). Despite the application of SPSA as a solution algorithm in numerous studies (Ma et al., 2007; Paz et al., 2012;

Vaze et al., 2009), the deterioration of performance with growing size of network and simulation period motivates the development of a better algorithm. Lu (2013) proposed a weighted SPSA (or W-SPSA) algorithm combining the principles of SPSA with a more reliable gradient approximation method, which outperforms SPSA for large-scale networks.

5.2.2   Calibration with multisource traffic data

Growing use of AVI technologies provides useful point-to-point flow information for estimating network OD matrix, and travel time data is a typical example. When formulating the typical objective function for OD estimation problem, an additional term is added to consider the deviation of simulated link travel time from observed link travel time extracted from AVI data (Djukic et al., 2015; Vaze et al., 2009).

However, this approach does not fully utilize the useful information of OD distribution which can be inferred from AVI data. One main challenge is to properly relate AVI data with OD distribution. Antoniou et al. (2006) applied Kalman filtering to solve the OD estimation problem based on state-space modeling. A diagonal scaling matrix is assumed as given input to capture the market penetration pattern of the probe vehicles in the total vehicle population. However, this study did not address how to obtain the scaling matrix (or market penetration rate matrix).

To further address the market penetration issue, Van Aerde et al. (1993) proposed an approach to account for both the temporal and spatial variation of market penetration rates. While temporal variation is captured by defining market penetration rate for a certain period t, spatial variation is considered by calculating penetration rate for each OD pair. Then the estimate of demand for a given OD is calculated by number of probe vehicles captured for that OD and the estimated penetration rate. Asakura et al. (2000) proposed a similar approach to estimate the percentage of true OD flow observed with the AVI system, and developed the relationship between population OD flow and observed OD flow in a linear equation.

Due to potential bias and difficulties in estimating market penetration rate, some

studies take a different approach for the use of AVI data. Zhou and Mahmassani (2006) proposed a new OD estimation method with AVI data, without the need to estimate market penetration rate. The objective function for OD estimation combines deviation for link counts, historical static demand and AVI split fractions. The formulation is based on the assumption that AVI tagged vehicles are a representative subset of the entire population. Similarly, Dixon and Rilett (2002) extracted the OD split proportion, choice proportion and travel times from AVI data as measurement from AVI data, and incorporated them with link count data to conduct OD estimation. AVI penetration rate is not considered in the study.

5.2.3   Calibration of close-loop framework

As we are using MITSIM to mimic the real-world traffic system, calibration of both DynaMIT and MITSIM are required before running the simulation. Depending on the role of each simulator, the traffic data as input for calibration is different. Figure 5.11 illustrates the best procedure in which in the first step the real-world traffic surveillance data is used to calibrate parameters in MITSIM. Then MITSIM would run with the calibrated parameters and output the simulated traffic data. The MITSIM simulated observation would be provided as input for DynaMIT calibration in the second step. The main idea is to replace the real-world traffic system with MITSIM, and use MITSIM to evaluate different control strategies generated by DynaMIT.
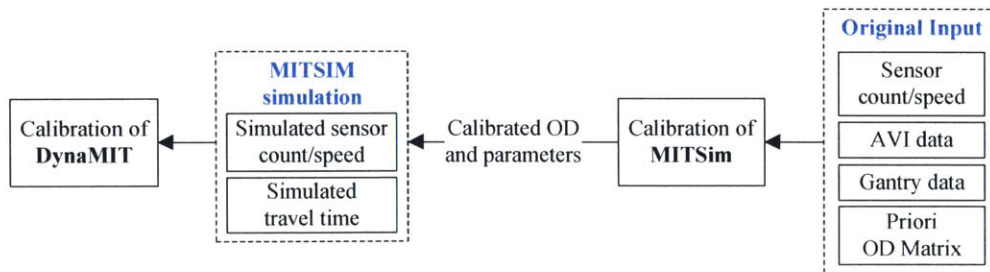

Figure 5.11 Calibration for close-loop framework

5.3 Calibration approach for NTE network

In this case study, the main complexity of calibration work is how to make efficient and effect use of AVI data in MITSIM **offline** calibration, and account for the difference for two types of drivers. The network and data for our case have the following characteristics:

(1) The AVI detectors are installed at all of the origin and destination nodes (i.e. at all the entrance and exit ramps to both general purpose lane and managed lane). Therefore, a full OD matrix for vehicle equipped with transponders can be constructed.

(2) For a number of the origin/destination nodes, there are also traditional loop detectors providing link count information. So it is possible to directly calculate the market penetration rates for these nodes, and then apply them to each OD pair through algorithms like bi-proportional apportionment.

(3) For some nodes, direct measurement of sensor count is not available. However, sensor count can be inferred from sensors located at upstream and downstream segments on main road.

(4) For origin/destination nodes with no market penetration rate, it can be assumed that traffic entering/exiting the same road segment group has similar market penetration patterns, and the average of penetration rate for all nodes in the particular segment group can be used to as an approximation.

(5) To further account for the potential bias and error in constructing OD matrix from AVI data and expanding it to population OD, the expanded OD is only used as seed values for further calibration.
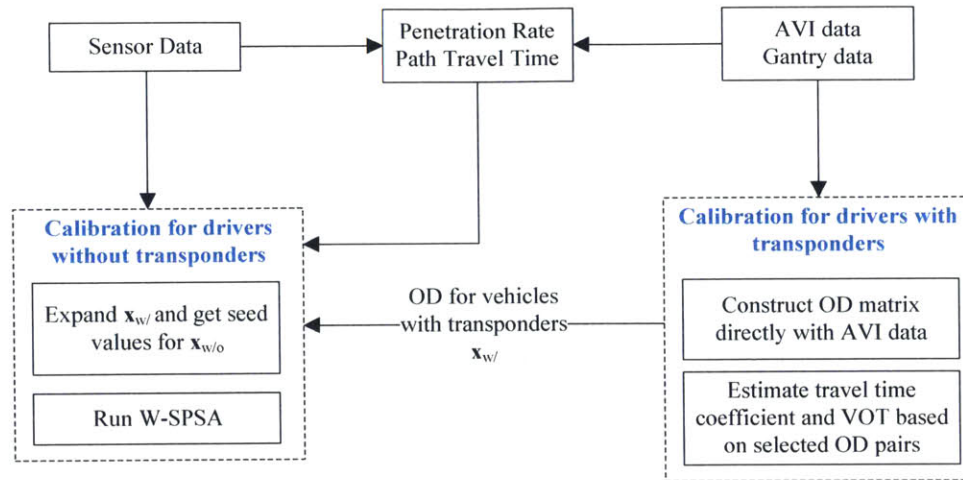
Figure 5.12 Procedure of offline calibration with multi-source traffic data

The calibration procedure is illustrated in Figure 5.12, and detailed steps are described as below:

Define group of drivers

(1)   Drivers with transponders (w/)

(2)   Drivers without transponders (w/o)

Define generic/group-specific parameters

Table 5.4 shows the parameters to be calibrated for two groups of drivers.

Table 5.4 Parameters to be calibrated for two groups of drivers

Parameters	DynaMIT	MITSIM	Drivers w/ transponders	Drivers w/o transponders
Time-dependent OD matrix	√	√	$x_{w/}$	$x_{w/o}$
Route choice parameters: a) Travel time coefficient, b) VOT distribution	√	√	$\beta_{RC,w/}$	$\beta_{RC,w/o}$
Driving behavior parameters	NA	√	β_{DB}	
Supply parameters: Speed-density function, etc.	√	NA	β_{SP}	

Calibration for drivers with transponders

(1) Construct $x_{w/}$ directly with AVI data.

(2) Select a set of major OD pairs with long enough travel time to differentiate managed and general purpose lanes, and where there are AVI detectors available to capture and distinguish drivers choosing managed and general purpose lanes. See Figure 5.13 for an illustration, where the black path on managed lane and grey path on general purpose lane are two comparable routes with the same OD.
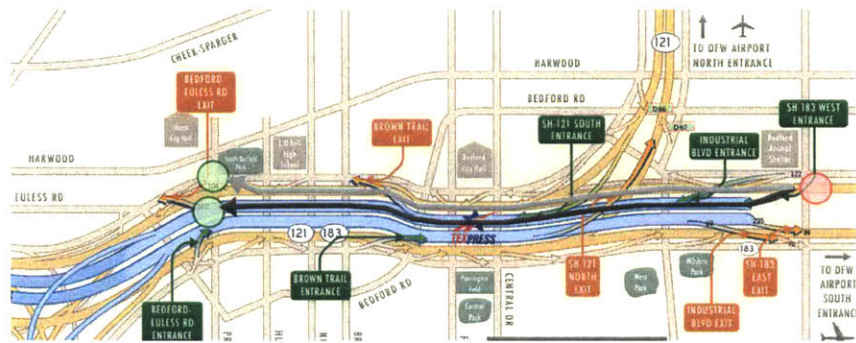
Figure 5.13 Example of selected OD pair

(3) Extract travel time and travel cost information for each driver and for each path. For each driver, extract *tt.ML* (travel time using managed lane), *tt.GP* (travel time using general purpose lane), *cost.ML* (monetary cost for paying the toll to use managed lane), *cost.GP* (0), and observed *path choice* (use managed lane or general purpose lane). By developing the binomial logit model, travel time coefficient and VOT can be estimated for vehicles with transponders.

Calibration for drivers w/o transponders

(1) Expand $x_{w/}$ to get seed values for $x_{w/o}$

Figure 5.14 illustrates the expansion of the OD matrix with two origin and two destination nodes. In Figure 5.15, the cells highlighted in gray in the table above represent the base OD obtained from AVI data, while the

cells in gray in the table below represent the sensor counts collected at each node. One possible method is to use Iterative Proportional Fitting (IPF) to obtain estimates of $POP_{O_i \to D_j}$ in cells highlighted in yellow so as to satisfy the sensor count constraints and maintain the seed distribution of $AVI_{O_i \to D_j}$.
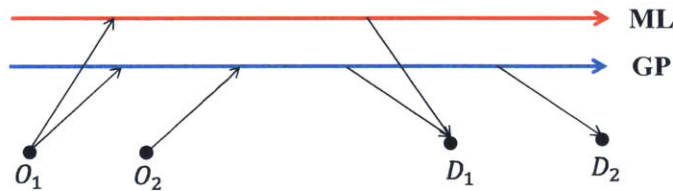
Figure 5.14 Network for illustration

	D_1	D_2	Total AVI vehs
O_1	$AVI_{O_1 \to D_1}$	$AVI_{O_1 \to D_2}$	AVI_{O_1}
O_2	$AVI_{O_2 \to D_1}$	$AVI_{O_2 \to D_2}$	AVI_{O_2}
Total AVI vehs	AVI_{D_1}	AVI_{D_2}	

	D_1	D_2	Sensor count
O_1	$POP_{O_1 \to D_1}$	$POP_{O_1 \to D_2}$	f_{O_1}
O_2	$POP_{O_2 \to D_1}$	$POP_{O_2 \to D_2}$	f_{O_2}
Sensor count	f_{D_1}	f_{D_2}	

Figure 5.15 Illustration of OD expansion using AVI and sensor data

(2) Extract link travel time

Select a set of major paths in the network, and extract time-dependent travel time from AVI data. In DynaMIT/MITSIM, we can add up simulated link travel time to get path travel time.

(3) Define objective function

The offline calibration is formulated as optimization problem 5.1. Constraint 5.2 shows the relationship between simulated outputs and model inputs. Constraint 5.3 gives the upper and lower bound of parameter values.

$$\min Z(\mathbf{x}, \boldsymbol{\beta}) = \sum_{h=1}^{H} [z_1(\mathbf{M}_h^o, \mathbf{M}_h^s) + z_2(\mathbf{x}_h, \mathbf{x}_h^a)] + z_3(\boldsymbol{\beta}, \boldsymbol{\beta}^a) \qquad 5.1$$

subject to:

$$\mathbf{M}_h^s = f(\mathbf{x}_h, \boldsymbol{\beta}, \mathbf{G}) \qquad 5.2$$

$$(l_{\mathbf{x}_h}, l_{\boldsymbol{\beta}}) \le (\mathbf{x}_h, \boldsymbol{\beta}) \le (u_{\mathbf{x}_h}, u_{\boldsymbol{\beta}}) \qquad 5.3$$

where G is the road network; x is the time dependent OD matrix; β includes other model parameters; M^o and M^s are the observed and simulated time-dependent traffic measurement, respectively (i.e. sensor count, sensor speed, travel time); x^a and β^a are priori values for timed dependent OD matrix and other model parameters, respectively. The goodness-of-fit measures z_1, z_2 and z_3 capture the error between the simulated or estimated quantities and their measured or a priori values.

Here component z_1 consists of three terms: the deviations of simulated sensor counts C, sensor speed V and link travel time L from real observation. The variables to be optimized are $x_{w/o}$ and $\beta_{RC,w/o}$, with $x_{w/}$ and $\beta_{RC,w/}$ being known inputs from previous steps. More detailed specification of z_1, z_2 and z_3 refers to (Ben-Akiva et al., 2010a; Lu, 2013).

While the general approach for DTA calibration considers the discrepancy between estimated parameters and historical values, in our study, high-quality historical values for behavior model parameters and OD matrix are not available, and the provided historical values are only used as starting points for calibration. Therefore, only deviation between simulated and observed measurements are considered for calibration.

(4)   Solution algorithm

W-SPSA (Weighted - Simultaneous Perturbation Stochastic Approximation

is used as solution algorithm, see (Lu, 2013) for details. The way SPSA estimates gradients would result in performance deterioration for a real-world traffic system in that all the parameters are perturbed at the same time no matter if the change in measurement values is caused by this parameter or not. A weighted gradient approximation method is used to exclude the influence from irrelevant measurement when approximating the gradient, and measurements are weighed based on their correlation with a specific parameter. After running N iterations of W-SPSA, further perturbation will switch to SPSA. Therefore, the main advantage of W-SPSA compared with traditional SPSA is its capability of reducing the noise generated by irrelevant measurements where correlations between parameters and measurements are sparse.

(5)   Goodness of fit:

Normalized Root Mean Square Error (RMSN) is used as a typical measurement of goodness of fit:

$$RMSN = \frac{\sqrt{N\sum_{i=1}^{N}(M_i^o - M_i^s)^2}}{\sum_{i=1}^{N}M_i} \qquad\qquad 5.4$$

where N is the number of observations; M_i^o and M_i^s are observed and simulated values for the ith measurement, respectively.

5.4 Calibration results

Due to the limitation of DynaMIT and MITSIM in simulating different types of vehicles with their own route choice models, in this section, initial calibration results with only sensor data are demonstrated. The simulation period is from 5:30 to 9:00 AM on June 8, 2015. The first half an hour (i.e. 5:30-6:00 AM) is set as warm-up

period for which the deviation of sensor counts and speed are not incorporated in the output performance measure. The time interval for OD and sensor counts is set as 5 minutes.

MITSIM calibration

Two groups of parameters are to be calibrated for the case study network: **driving behavior** parameters and **travel behavior** parameters.

Driving behavior models mainly include car-following model and lane changing model. Seed values for these model parameters are obtained from previous studies. For general methodologies and typical model values of MITSIM calibration, we refer to (Ben-Akiva et al., 2000; Ben-Akiva et al., 2002b; Darda, 2002).

Travel behavior parameters include route choice model parameters and time-dependent OD matrix. A static planning OD matrix for a typical weekday is provided by road operator, and then used to generate an seed time-dependent OD matrix. Specifically, the original planning OD matrix provides the typical daily demand for each OD pair, and with the help of sensor data, we can calculate the total network sensor counts summed for each time period (i.e. every 5 minutes). Then the daily OD matrix can be converted to time-dependent OD matrix to match the sensor count variation throughout the day. This time-dependent OD matrix is used as seed value for further calibration.

Figure 5.16 shows the observed sensor count versus the MITSIM simulated sensor count with seed OD matrix and parameter values. Obviously, there is significant deviation between the two sets of traffic flow data. After running W-SPSA calibration codes, as shown in Figure 5.17 and Figure 5.18, most of the points are close to the 45-degree line, which indicates that the simulated sensor counts and speed are close to the observed sensor counts and speed.
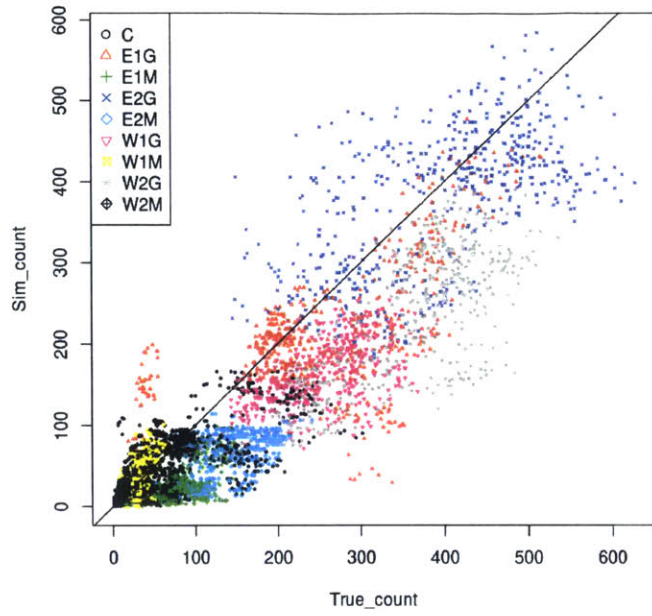
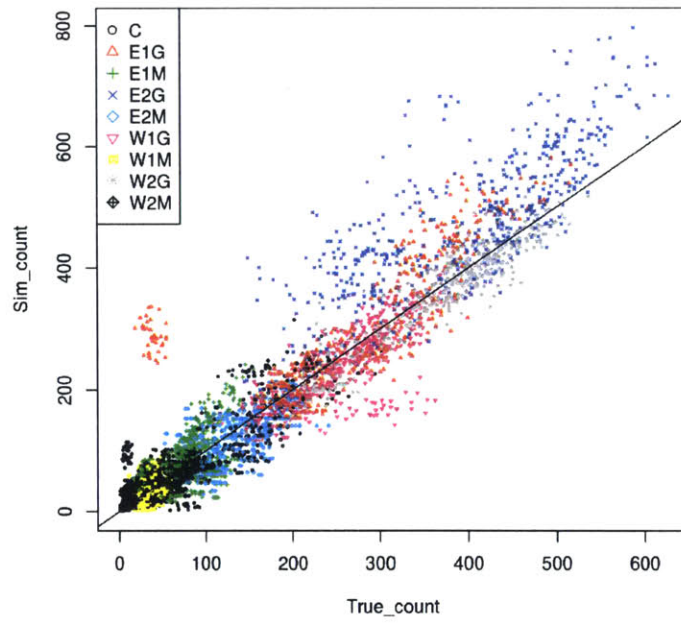Figure 5.16 Observed count versus simulated count (5min interval) at first run

Figure 5.17 Observed count versus simulated count (5min interval) at last run

73
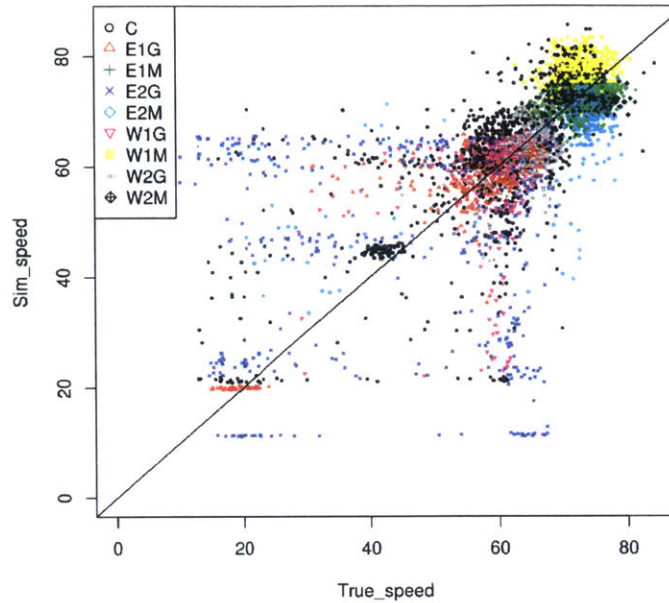
Figure 5.18 Observed speed versus simulated speed (5min interval) at last run

DynaMIT calibration

After MITSIM is calibrated, the simulated sensor flow and speed are used as inputs for DynaMIT calibration. On DynaMIT side, the parameter can be categorized by **demand** and **supply** parameters.

On demand side, the MITSIM calibrated OD matrix and route choice model parameters are used as the starting values for DynaMIT calibration.

On supply side, the parameters for speed-density relationship and segment capacities are first fitted by field data using typical curve fitting methods. Note that since not all segments have sensors on them, the sensor data are merged by segment type. Figure 5.19 illustrates the speed-density plot for a particular road segment group, and the red line represents the fitted curve. The fitted parameters will then further be refined, jointly with demand parameters, by W-SPSA calibration codes.

Figure 5.20 illustrates the comparison between MITSIM simulated sensor count and DynaMIT simulated sensor count, which generally shows good fitting.

Figure 5.19 Speed-density plot for a given segment group

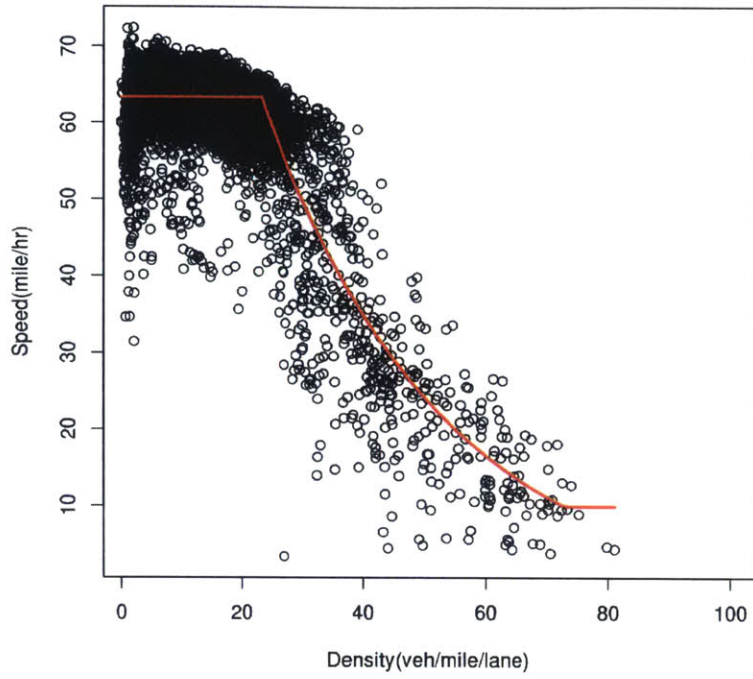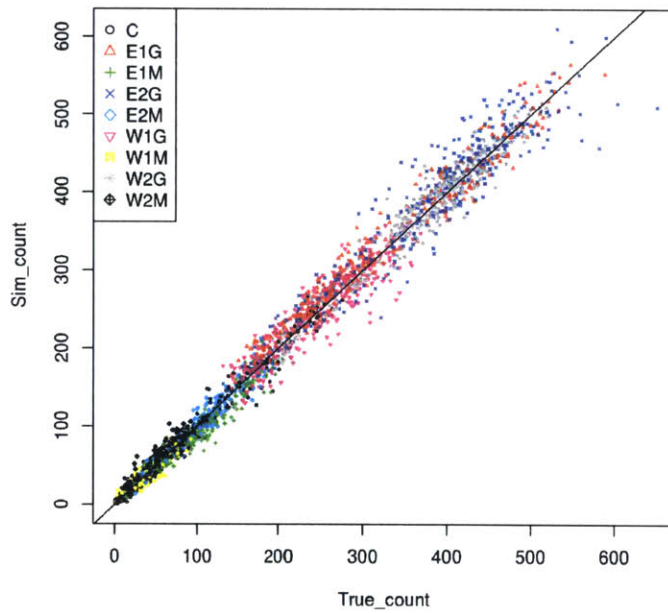Figure 5.20 Observed count versus simulated count

Goodness-of-fit

Table 5.5 listed the goodness-of-fit for both MITSIM and DynaMIT offline

75

calibration. The error rates are generally within the acceptable range, which indicates that the offline calibration well replicated the typical real-life traffic condition on this network in the two simulators.

Table 5.5 Goodness-of-fit for calibration

MITSIM	
RMSN Count	32.8%
RMSN Speed	17.2%
DynaMIT	
RMSN Count	9.8%
RMSN Speed	16.0%

5.5 Tolling Regulations

In the NTE TEXpress network, the road operator has to comply with contractual requirements, i.e., regulations defined by Texas Department of Transportation when imposing dynamic tolling. This section first illustrates the basic principles of regulation, and proposes an extended formulation of the optimization model under the tolling regulations. The aim is to demonstrate the capability of our framework to take into account regulation constraints in industry.

5.5.1   Overview of regulation rules

(1) *Segment toll* is defined as the *base toll rate* multiplied by the toll segment length.

(2) *Base toll rate cap* is the upper bound of base toll rate (in dollar) per mile per *toll segment direction*.

(3) The road operator cannot change the base toll rate more frequently than every five minutes.

(4) If the road operator wants to set the base toll rate higher the base toll cap rate for a toll segment direction, toll setting mode must enter *mandatory mode*. When in mandatory mode, the road operator should comply with the following provisions:

76

- At the end of each tolling interval, the *measured average speed* $\bar{v}$ and *measured average volume* $\bar{q}$ are calculated by aggregating all sensors located at the toll segment direction.

- If $\bar{v}$ is lower than v_{cr}, base toll rate shall be multiplied by a flexible *demand factor* between a lower bound DF_{lb} and a higher bound DF_{ub}, and toll rate will increase compared with previous toll (i.e. $DF_{lb} \geq 1$).

- If $\bar{v}$ is higher than v_{cr}, depending on the level $\bar{q}$, there is a set of pre-defined rules to calculate a fixed demand factor DF. The base toll *rate* shall be multiplied by DF, which may increase, decrease and maintain the previous toll.

5.5.2 Mathematical formulation

The aim is to incorporate the regulations into the current framework, so that the decision on whether to switch to mandatory mode or not is determined by field surveillance data, while the optimizer should be capable of modelling and predicting the decisions for future intervals based on the forecasted traffic condition.

Table 5.6 Additional notation for incorporating regulation rules

η_{it}	binary variable, whether we are allowed to enter mandatory mode or not.
δ_{it}	binary variable, whether we decide to enter mandatory mode or not.
DF_{lb}^{it}, DF_{ub}^{it}	lower and upper bounds for the demand factor

Here η_{it} is input for the next interval based on field measurements (provided by MITSIM in the closed-loop setting) and a variable to be optimized for the subsequent intervals based on predicted traffic state. Similarly, DF_{lb}^{it} and DF_{ub}^{it} are inputs for the immediate next interval and variables for the subsequent intervals

Also, additional constraints are introduced to model the entering/existing decisions on mandatory mode, and detailed interpretation of these constraints is explained.

$$(v_{it}, q_{it}) = DynaMIT(\beta) \qquad \forall i \in I, t \in T \qquad\qquad 5.5$$

$$\delta_{it} \leq \eta_{it} \qquad\qquad \forall i \in I, t \in T \qquad\qquad 5.6$$

$$\delta_{it} \geq M(\delta_{i(t-1)} - 1) + (1/100)(\beta_{i(t-1)} - \beta_{CAP}) \qquad \forall i \in I, t \in T \qquad 5.7$$

$$(DF_{it}^{lb}, DF_{it}^{ub}, \eta_{it}) = f(v_{it}, q_{it}) \qquad \forall i \in I, t \in \{2,...,T\} \qquad 5.8$$

$$\delta_{it}\beta_{i(t-1)}DF_{it}^{lb} \leq \beta_{it} \leq (1-\delta_{it})\beta_{CAP} + \delta_{it}\beta_{i(t-1)}DF_{it}^{ub} \quad \forall i \in I, t \in T \qquad 5.9$$

$$\beta_{i(t-1)} - \Delta - \delta_{it}M \leq \beta_{it} \leq \beta_{i(t-1)} + \Delta + \delta_{it}M \qquad \forall i \in I, t \in T \qquad 5.10$$

Equation 5.5: Predicted speed and volume are provided by traffic simulator DynaMIT to evaluate objective function and also for the decisions in future intervals.

Equation 5.6: System cannot enter mandatory mode if not allowed by measurements (for the next interval) or predictions (for future intervals).

Equation 5.7: If system was in mandatory mode in t-1 and if the toll was above the toll cap, system needs to stay in mandatory mode. For example, if the traffic conditions are getting better, this enables that toll is reduced gradually based on regulations.

Equation 5.8: Functions are predefined by road operator to get the lower and upper bounds on the demand factor and the allowance to enter mandatory mode for the future intervals based on predictions. Note that DF and η values are inputs for the first interval and variables for the future intervals.

Equation 5.9: If the decision is to stay in dynamic mode (δ=0), then the toll is optimized between 0 and toll cap, otherwise (δ=1) toll rates follow the regulations in mandatory mode.

Equation 5.10: Maximum change in the toll is imposed, and this constraint is active only in dynamic mode (δ=0), but will not be binding in mandatory mode (δ=1)

5.5.3 Testing on synthetic network

Initial results are demonstrated with the synthetic network presented in Chapter 4. Demand profile is specified to have an increasing trend at 6:30-7:15 AM and decreasing trend at 7:15-8:00 AM. The aim of this setting is to capture cases to enter and exit the mandatory mode. Note that for this setting, toll cap is calculated as $0.95 based on segment length.

Figure 5.21 illustrates the optimal toll rates, where the system is capable of reacting to the changing traffic demand input. System first decides to enter mandatory mode because of the degrading traffic conditions on managed lane. Afterwards, due to the decrease in traffic demand, the system is no longer allowed to stay in mandatory mode, so it starts to gradually decrease toll rate so as to exit mandatory mode.

Figure 5.21 Optimal tolls with regulation rules
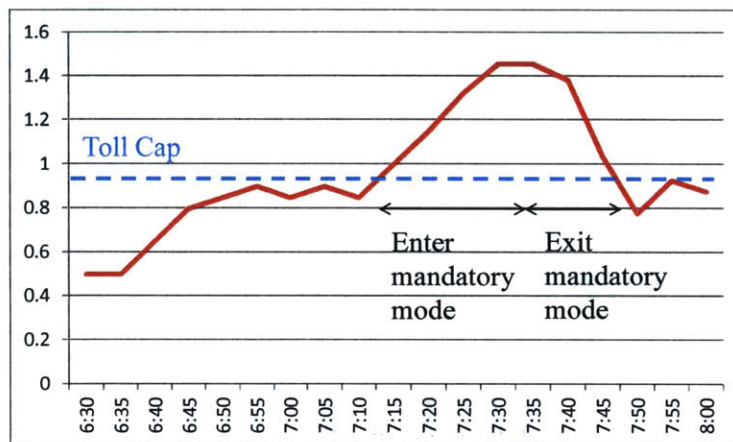
Table 5.7 lists the decision for each interval, which corresponds to Figure 5.21. When the detected flow goes beyond the predefined threshold (highlighted in red), the system is allowed to enter mandatory mode. On the contrary, when demand is decreasing and congestion starts to relieve (highlighted in green), system is forced to decrease the toll rates until exiting mandatory mode.

Table 5.7 Optimal tolls with regulation rules

Time	η	DF	speed (mph)	flow (veh/2 lanes)	β ($)	δ
6:30			warm-up			
6:35						
6:40	0	---	59.00	1536	0.65	0
6:45	0	---	59.00	1656	0.80	0
6:50	0	---	58.55	1668	0.85	0
6:55	0	---	58.15	1950	0.90	0
7:00	0	---	55.70	2526	0.85	0
7:05	0	---	53.35	2952	0.90	0
7:10	0	---	52.95	3138	0.85	0
7:15	1	1.1	51.75	3492	1.00	1
7:20	1	1.15	53.10	3552	1.15	1
7:25	1	1.15	54.35	3522	1.32	1
7:30	1	1.1	51.30	3408	1.45	1
7:35	1	1	50.10	2844	1.45	1
7:40	1	0.95	52.45	2436	1.38	1
7:45	1	0.75	56.80	1890	1.04	1
7:50	1	0.75	59.65	1584	0.78	1
7:55	0	---	59.80	1614	0.93	0
8:00	0	---	59.75	1662	0.88	0

5.6 Optimization results

Preliminary optimization is performed with calibrated OD parameters using only sensor data as described in section 5.1.3. The simulation period is 5:30-9:00 AM, with the first half an hour set as warm-up period. The starting toll rates for each segment at 5:30 AM is set as $2.50, $4.50, $2.50, $2.50 for segment group S1E, S2E, S1W, S2W, respectively based on historical averaged toll rates. Tolling interval is 5 minutes, and prediction horizon is 15 minutes. Exhaustive search is utilized for generating the results. At each interval and for each toll segment, the system will search for the +$0.10 and -$0.10 toll change compared to the previous interval. All the possible toll value combinations for the four toll segments are evaluated using the predicted revenue over the 15-minute prediction horizon as the objective function. The toll vector with maximal predicted revenue in DynaMIT prediction is selected as the

optimal control strategy to be sent to MITSIM.

The evaluation is performed in parallel by running the state prediction module of DynaMIT independently for each toll rate input. The experiment is operated on a computer equipped with Intel Core i7 CPU and eight cores. So eight toll rates can be evaluated in parallel. The computation time for running the whole simulation period with toll optimization is around 15 minutes measured in amount of CPU time.

Figure 5.22 (a) illustrates the optimized toll rates for each toll segment. The optimized toll rates for S2E are higher. For other segments, the toll rates are relatively stable and slightly increasing as time evolves. Compared to the base tolls in Figure 5.22 (b), which are the toll rates implemented and recorded by road operator on the same day, the general pattern is similar, except that the base tolls are higher than optimized tolls for all segments.

Here there are three cases for further comparison of traffic conditions

(1) "Optimized toll" case is simulated in the close-loop environment where MITSIM received predictive travel time guidance and optimal toll rates from DynaMIT every time interval.

(2) "Base toll (simulated)" case is simulated using MITSIM without guidance from DynaMIT. The input toll rates are the values recorded by road operator during the same time period. The simulated results are the same as the outputs from off-line calibration.

(3) "Base toll (observed)" case is the observed traffic condition from field data.

Figure 5.23 (a), (b) and (c) compare the flow at gantries under different cases. Figure 5.24 (a), (b) and (c) compare the simulated sensor speed aggregated by segment group under different cases. There is some difference in traffic conditions among three cases. But generally, the speed on managed lane is always maintained at free flow speed. The flows under the three cases are generally consistent with the toll rates, such that the number of vehicles passing the gantry decreases as the toll grows.
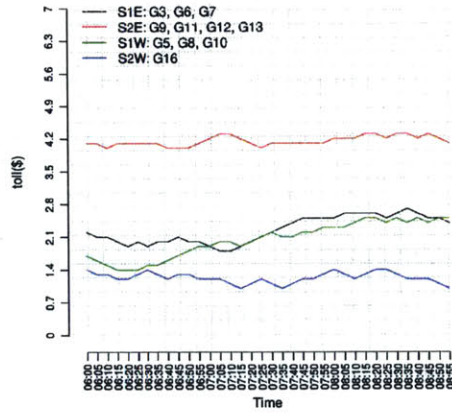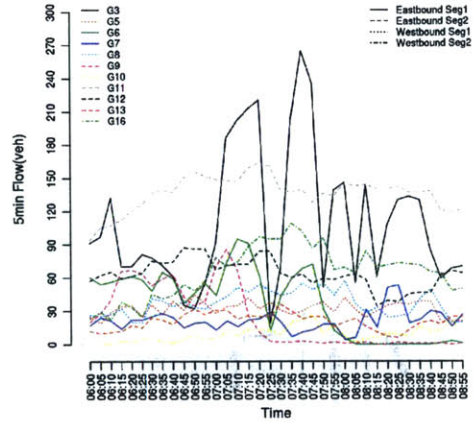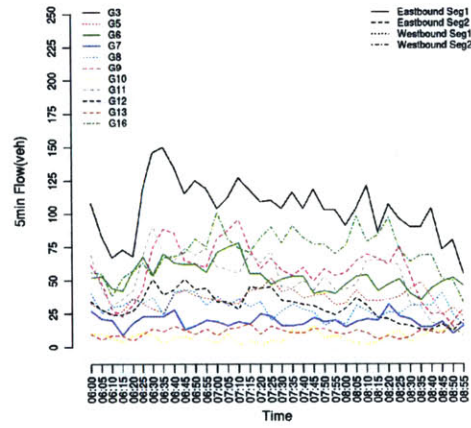
(a) Optimized toll
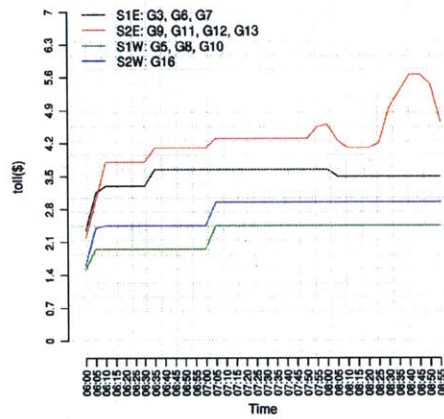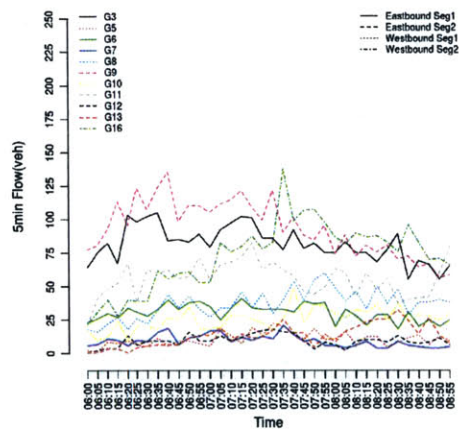

(b) Base toll

Figure 5.22 Comparison between optimized tolls and base tolls
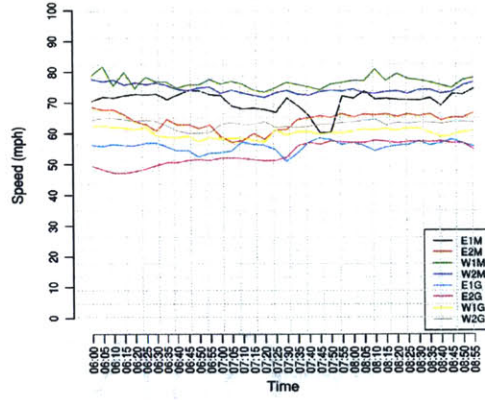
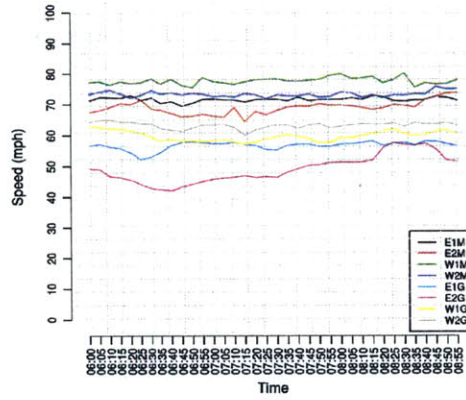(a) Optimized toll

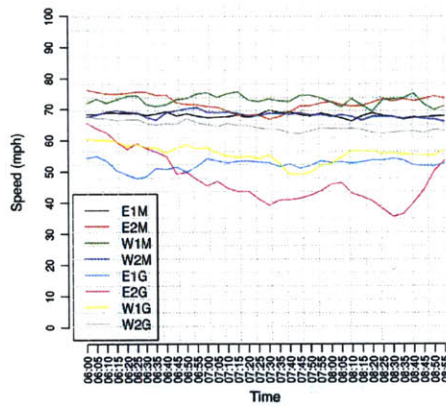(b) Base toll (simulated)

(c) Base toll (observed)

Figure 5.23 Comparison of flow at gantries

83

(a) Optimized toll

(b) Base toll (simulated)

(c) Base toll (observed)

Figure 5.24 Comparison of average sensor speed

To further compare the optimal toll strategy against the benchmark case, Figure 5.25 illustrates the comparison between optimized toll and base toll in simulated revenue from MITSIM over time. The total observed revenue on field is $48612. The total simulated revenue under calibration is $46186 with base toll rates. The total simulated revenue is $48661 with optimized toll rates. Overall, the revenue values provide a good match of the real-life conditions, and there is slight improvement.
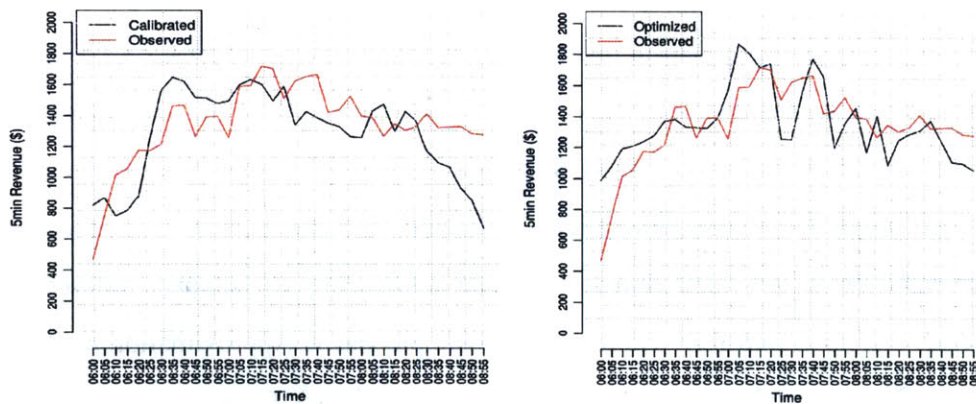
(a) Calibrated vs. Observed   (b) Optimized vs. Observed

Figure 5.25 Comparison of simulated revenue

These simulation outputs, analysis and comparison are preliminary and still need to be further investigated for the following reasons:

(1) Possible explanations for lower toll rates compared with base toll may be insufficient calibration of drivers' route choice model so that our model underestimates driver's willingness to pay for using the managed lane.

(2) There might be different assumptions in drivers' access to travel time guidance information. Also, in reality drivers may make their decision based on additional considerations, like travel time reliability or vehicle type discount.

(3) The results need to be further validated with the field observed data to check if the simulation captures the network bottleneck, congestion point, etc.

5.7 Considerations on solution algorithm

Currently the preliminary results are based on exhaustive search, which requires considerable computation time to search for the whole solution space. If the optimization problem becomes more complicated, computational efficiency would be a critical issue to address if the system is to be operated in real time manner. For example, the growing number of tolling location, the increase in length of prediction horizon and the larger network scale will all increase computational burden.

The optimization problem in this study is essentially to optimize for a black-box function, since the value of objective function is evaluated by traffic simulator, and there is no closed form for the output. When the framework is applied to a large-scale network, computation time and efficiency should be addressed so that it can be implemented in a real time manner. Typical approaches to solving optimization problems with no closed-form expression include Stochastic Approximation methods, Response Surface Methodology (Surrogated-based methods), Genetic Algorithm, etc. Here genetic algorithm and surrogate-based methods will be discussed.

Genetic algorithm is preferable to solve problem with multiple local optima and a large number of parameters (Q. Wang, 1997), which outperforms traditional gradient based methods. Genetic algorithm has been previously used in solving transportation problems like OD estimation (Zou and Kulkarni, 2013). Gupta et al. (2016) applied genetic algorithm in solving dynamic road pricing problem as well. Genetic algorithm with parallel evaluation of population using parallel and distributed computing techniques was implemented in developing real time tolling strategy.

This algorithm will start by randomly generating a group of control strategies (i.e. toll vectors) as initial individuals. These individuals will form the initial populations or parent population with a predefined size N. Then these N individuals will be evaluated in traffic simulation in a parallel and independent way, and the output would be objective function value for each of the individual. Based on the ranks of the objective function values, a new set of toll vectors which are called child population, will be generated using a genetic operator. Typically crossover and mutation between

different individuals will be conducted in this process. The updated generation of population will be evaluated again in traffic simulator. The procedure of generating new child population will be repeated until termination criteria is satisfied, which can be a predefined maximum number of iterations or a threshold of improvement between consecutive generations.

Surrogate-based optimization represents a class of optimization methodology, which utilizes surrogate modelling techniques to search for the optima. It is referred to as a technique that makes use of the sampled data (observed by running the simulation model) to build surrogate models, which are sufficient to predict the output of an expensive simulation run at untried points in the design space (Han and Zhang, 2012). Surrogate modelling has been recently applied in simulation-based optimization for road pricing (X. M. Chen et al., 2014) and signal timing (Osorio and Chong, 2015). Figure 5.26 illustrates the general procedure of using surrogated-based methods if DynaMIT state prediction is to be utilized as the high-fidelity model for evaluating each sample point.
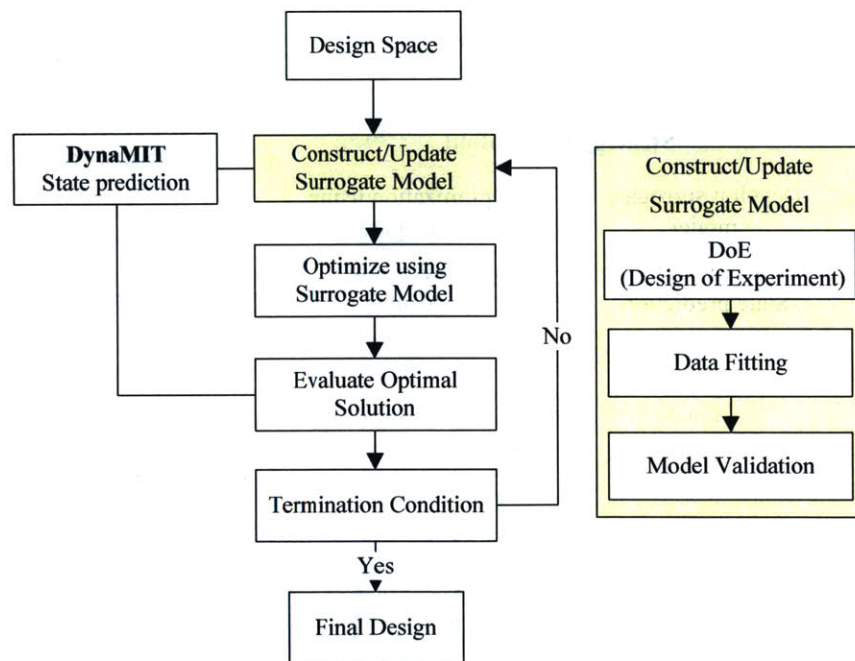
Figure 5.26 Illustration of possible framework using surrogate models

Figure 5.27 further illustrates a possible way of utilizing surrogate models in the real time optimization framework in this study. The potential capability of computational time saving and its tradeoff with optimality still need to be further investigated. Moreover, if surrogate models are to be built in real time, how to intelligently use surrogate model and/or optimal toll rates from last interval to optimize for the current interval can be interesting to further consider, since the information from previous intervals may improve the initial sampling, and reduce the design space for optimal solution. It might be possible to build surrogate models offline based on historical database with different types of network state and demand, so that the online efforts of building surrogate models for each interval may be relieved.

Figure 5.27 Initial design with surrogate-based optimization

6 Conclusion

6.1 Summary

Road pricing, especially pricing on managed lanes, which is the focus of this thesis, has a variety of aspects interesting for research. The framework of developing road pricing strategy is also crucial for assessing the feasibility and effectiveness of a managed lane project. Nowadays, communication and vehicle detection technology has gifted planners and road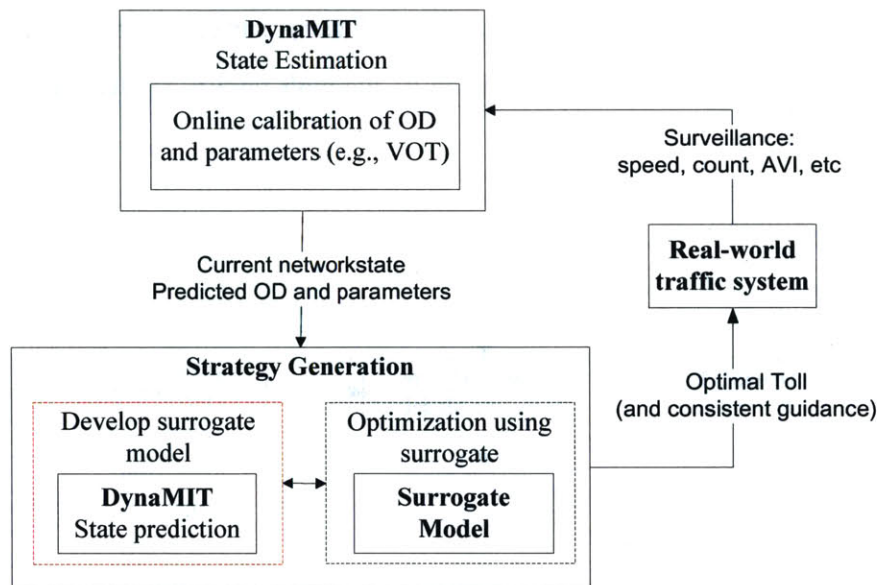 operators with the ability to develop more efficient and flexible tolling strategies to accommodate for the fluctuating traffic demand and changing network states.

Extensive study on existing managed lane facilities reveals that most road operators choose time-of-day strategy or dynamic tolling strategy to account for the different demand levels throughout the day. The rules of toll updating, however, mainly based on empirical data or professional intuition. Academic research on road pricing mainly adopts reactive control, proactive control, optimization-based control. The tool used for modelling traffic flow dynamics also vary from simplified, analytical, macroscopic models to sophisticated, stochastic, microscopic models.

This thesis proposes a real time optimization framework for dynamic pricing on managed lanes. The main characteristics of the proposed framework are:

(1) **Real time**: The system runs on a rolling-horizon scheme. Each time interval traffic control center, DynaMIT, will receive traffic surveillance data, run the optimization and return the optimal toll values to be implemented for the immediate next interval.

(2) **Prediction-based**: The toll rates are evaluated based on the predicted traffic condition and drivers' travel behavior. Therefore, for each possible solution, we can forecast the impact of such a control strategy for network performance and driver's behavior.

(3) **Consistent**: The system will generate guidance information for drivers in the traffic system, and this guidance information is consistent in the sense that,

drivers receiving the information on road will experience a similar travel time given the path they have chosen and the toll charged on them.

(4) **Adaptive**: The system will calibrate the model parameters and OD matrix embedded in the simulator in a real time manner, so that the latest field data is utilized by the system to adjust the network state and driver behavior models.

Mathematical formulation of the optimization problem is proposed, and a close-loop testing framework is developed which relies on the interaction between the toll optimization model, the prediction capabilities of DynaMIT and the microscopic simulation of MITSIM. The framework is first analyzed with numerical experiments on a small example network with different formulations and behavioral assumptions. The results demonstrated the capability of the system reacting to changing demand and behavior models.

Then a case study on NTE TEXpress Lanes network is conducted to apply the optimization framework on a real-life highway network. A number of practical issues are addressed for the network, including field data filtering, extracting information for vehicles equipped with transponders, study on the current tolling strategy, etc. For calibrating the network and making better use of available data, a calibration approach combining data from multiple sources is proposed. Initial calibration results with only sensor count and speed are presented and evaluated.

To take into account the regulation rules from local government on managed lane operation and constraints on tolling strategy, mathematical formulations are proposed for these rules so that the system can provide tolling plan while satisfying all the practical restrictions. Then preliminary results with optimized tolls are presented for morning peak hours on a typical weekday. Optimal toll rates for each interval, simulated network state and comparison with the benchmark toll strategy are shown in the analysis. The results demonstrated the capability of the framework applying on a large-scale network. The real time performance of the system is critical to be maintained for large-scale networks. A few promising solution methodologies are discussed and possible integration with the current framework is proposed.

6.2 Future research

6.2.1 Travel behavior model

Drivers' route choice between managed lane and general purpose lane is one fundamental part of the framework compared with urban network with no tolls. Currently the route choice model embedded in the DTA model considers the typical formulation of drivers' utility, and there is still room of improvement. Further study may investigate the impact of travel time reliability, group-specific distribution of model parameters, different specification for travel time/cost, etc.

Meanwhile, it would be interesting to extract more individual information for transponders, and drivers' attributes can be interacted with travel time and cost to better predict their route choice. For example, truck drivers may have significantly different choice models compared with automobile drivers because they may have their employers pay the tolls. Also, drivers' gender, age, residence location and other demographic attribute may also enter the choice model. More realistic modelling will include drivers' mode choice and departure time choice as well.

6.2.2 Efficient algorithm

One main difference of this study from previous study on toll optimization is that we aim to run optimization in a rolling horizon scheme instead of one-shot optimization at the beginning, so I am interested in extending our model to utilize advanced optimization methods, possibly revised surrogate model, to improve computational efficiency. Also, further research can be done to see how we can make use of historical database to improve the efficiency of optimal toll searching. Possible approach can be constructing a library of representative network states for a given period of time. And for each network state in this database, an offline optimization can be run to obtain the optimal toll rates for the next interval. Then during the rolling horizon optimization, by matching the past network states with the ones in the historical database, the search space may be reduced to the neighboring values of

offline optimized toll rates.

6.2.3   Online calibration

Due to the important role of route choice model and its parameters in determining drivers' choice between general purpose lane and managed lane, ideally the route choice model parameters, instead of calibrated only offline once a day or once a week, should be capable of changing dynamically to the latest surveillance data as well. Challenges for this function include how to make joint use of AVI data and sensor data in calibrating route choice parameters for every five minutes, how to deal with drivers in the simulators whose observed route choice is different from estimated, etc. Meanwhile, efficient solution algorithm should also be developed for online calibration so as to meet the real time optimization requirement.

Reference

Antoniou, C., Ben-Akiva, M., & Koutsopoulos, H. N. (2006). *Dynamic traffic demand prediction using conventional and emerging data sources.* Paper presented at the IEE Proceedings-Intelligent Transport Systems.

Asakura, Y., Hato, E., & Kashiwadani, M. (2000). Origin-destination matrices estimation model using automatic vehicle identification data and its application to the Han-Shin expressway network. *Transportation, 27*(4), 419-438.

Balakrishna, R., Ben-Akiva, M., & Koutsopoulos, H. (2007). Offline calibration of dynamic traffic assignment: simultaneous demand-and-supply estimation. *Transportation Research Record: Journal of the Transportation Research Board.*

Bar-Gera, H. (2012). *Fast Lane to Tel-Aviv: High-Occupancy-Toll Project with Pareto Package.* Paper presented at the Transportation Research Board 91st Annual Meeting.

Bekhor, S., Ben-Akiva, M., & Scott Ramming, M. (2002). Adaptation of logit kernel to route choice situation. *Transportation Research Record: Journal of the Transportation Research Board*(1805), 78-85.

Ben-Akiva, M., Bierlaire, M., Burton, D., Koutsopoulos, H. N., & Mishalani, R. (2001). Network state estimation and prediction for real-time traffic management. *Networks and Spatial Economics, 1*(3-4), 293-318.

Ben-Akiva, M., Bierlaire, M., Koutsopoulos, H. N., & Mishalani, R. (2002a). Real time simulation of traffic demand-supply interactions within DynaMIT *Transportation and network analysis: current trends* (pp. 19-36): Springer.

Ben-Akiva, M., Davol, A., Toledo, T., Koutsopoulos, H. N., Burghout, W., Andréasson, I., Johansson, T., & Lundin, C. (2000). MITSIMLab for Stockholm Enhancements, Calibration and Validation.

Ben-Akiva, M., Davol, A., Toledo, T., Koutsopoulos, H. N., Burghout, W., Andréasson, I., Johansson, T., & Lundin, C. (2002b). *Calibration and evaluation of MITSIMLab in Stockholm.* Paper presented at the 82nd Annual Meeting Preprint CDROM, Transportation Research Board.

Ben-Akiva, M., Koutsopoulos, H. N., Antoniou, C., & Balakrishna, R. (2010a). Traffic simulation with DynaMIT *Fundamentals of traffic simulation* (pp. 363-398): Springer.

Ben-Akiva, M., Koutsopoulos, H. N., Toledo, T., Yang, Q., Choudhury, C. F., Antoniou, C., & Balakrishna, R. (2010b). Traffic simulation with MITSIMLab *Fundamentals of Traffic*

Simulation (pp. 233-268): Springer.

Bottom, J. A. (2000). *Consistent anticipatory route guidance.* PhD thesis, Massachusetts Institute of Technology.

Boyles, S. D., Gardner, L. M., & Bar-Gera, H. (2015). Incorporating Departure Time Choice into High-occupancy/toll (HOT) Algorithm Evaluation. *Transportation Research Procedia, 9*, 90-105.

Buckeye, K. (2012). Performance Evaluation of I-394 MnPASS Express Lanes in Minnesota. *Transportation Research Record: Journal of the Transportation Research Board*(2278), 153-162.

Cambridge Systematics. (2006). Inc. I-394 MnPASS technical evaluation: Final report: Technical report, MnPASS, Minnesota Department of Transportation.

Chen, X., Osorio, C., Marsico, M., Talas, M., Gao, J., & Zhang, S. (2015). *Simulation-based Adaptive Traffic Signal Control Algorithm.* Paper presented at the Transportation Research Board 94th Annual Meeting.

Chen, X. M., Zhang, L., He, X., Xiong, C., & Li, Z. (2014). Surrogate‐Based Optimization of Expensive‐to‐Evaluate Objective for Optimal Highway Toll Charges in Transportation Network. *Computer‐Aided Civil and Infrastructure Engineering, 29*(5), 359-381.

Chung, C.-L. (2013). A Review and Advance of High-Occupancy Toll Lanes' Toll Schemes. *Journal of the Eastern Asia Society for Transportation Studies, 10*(0), 240-259.

Chung, C.-L., & Recker, W. (2011). *State-of-the-art assessment of toll rates for high-occupancy and toll lanes.* Paper presented at the Transportation Research Board 90th Annual Meeting.

Darda, D. (2002). *Joint calibration of a microscope traffic simulator and estimation of origin-destination flows.* Master thesis, Massachusetts Institute of Technology.

de Palma, A., & Lindsey, R. (2011). Traffic congestion pricing methodologies and technologies. *Transportation Research Part C: Emerging Technologies, 19*(6), 1377-1399.

Dixon, M. P., & Rilett, L. (2002). Real‐Time OD Estimation Using Automatic Vehicle Identification and Traffic Count Data. *Computer‐Aided Civil and Infrastructure Engineering, 17*(1), 7-21.

Djukic, T., Barceló, J., Bullejos, M., Montero Mercadé, L., Cipriani, E., van Lint, H., & Hoogendoorn, S. (2015). *Advanced traffic data for dynamic OD demand estimation: The*

state of the art and benchmark study. Paper presented at the TRB 94th Annual Meeting Compendium of Papers.

Dong, J., Mahmassani, H. S., Erdoğan, S., & Lu, C.-C. (2011). State-dependent pricing for real-time freeway management: Anticipatory versus reactive strategies. *Transportation Research Part C: Emerging Technologies, 19*(4), 644-657.

FHWA. (2008). *Managed Lanes: A Primer*. US Department of Transportation.

Florian, M., Mahut, M., & Tremblay, N. (2001). *A hybrid optimization-mesoscopic simulation dynamic traffic assignment model.* Paper presented at the Intelligent Transportation Systems, 2001. Proceedings. 2001 IEEE.

Göçmen, C., Phillips, R., & van Ryzin, G. (2015). *Revenue Maximizing Dynamic Tolls for Managed Lanes: A Simulation Study*. Columbia University Center for Pricing and Revenue Management.

Grzybowska, H., & Barceló, J. (2012). Decision support system for real-time urban freight management. *Procedia-Social and Behavioral Sciences, 39*, 712-725.

Gupta, S., Seshadri, R., Atasoy, B., Pereira, F. C., Wang, S., Vu, V.-A., Tan, G., Dong, W., Lu, Y., & Antoniou, C. (2016). *Real Time Optimization of Network Control Strategies in DynaMIT2. 0.* Paper presented at the Transportation Research Board 95th Annual Meeting.

Han, Z., & Zhang, K. (2012). *Surrogate-based optimization*, INTECH Open Access Publisher.

He, X., Chen, X., Xiong, C., Zhu, Z., & Zhang, L. (2016). Optimal Time-Varying Pricing for Toll Roads Under Multiple Objectives: A Simulation-Based Optimization Approach. *Transportation Science*. doi: http://dx.doi.org/10.1287/trsc.2015.0661

Ieromonachou, P., Potter, S., & Warren, J. P. (2006). Norway's urban toll rings: Evolving towards congestion charging? *Transport Policy, 13*(5), 367-378.

Jang, K., Chung, K., & Yeo, H. (2014). A dynamic pricing strategy for high occupancy toll lanes. *Transportation Research Part A: Policy and Practice, 67*, 69-80.

Joksimovic, D., Bliemer, M., & Bovy, P. (2005). *Dynamic road pricing optimization with heterogeneous users.* Paper presented at the 45th Congress of the European Regional Science Association, Amsterdam.

Li, J., & Govind, S. (2002). *An optimization model for assessing pricing strategies of managed lanes.* Paper presented at the Proc., 82nd Annual Meeting of the Transportation Research Board.

Liem, R. P. (2007). *Surrogate modeling for large-scale black-box systems.* Master thesis, Massachusetts Institute of Technology.

Lu, L. (2013). *W-SPSA: an Efficient Stochastic Approximation Algorithm for the off-line calibration of Dynamic Traffic Assignment models.* Master thesis, Massachusetts Institute of Technology.

Ma, J., Dong, H., & Zhang, H. (2007). Calibration of microsimulation with heuristic optimization methods. *Transportation Research Record: Journal of the Transportation Research Board.*

Mahmassani, H. S. (2001). Dynamic network traffic assignment and simulation methodology for advanced system management applications. *Networks and Spatial Economics, 1*(3-4), 267-292.

Martin, P. T., Vladisavljevic, I., Ries, J., & Nadimpalli, B. R. (2009). Express Lane Genetic Algorithm Microsimulation Evaluation Part (2): ELGAME 2: Mountain-Plains Consortium.

Michalaka, D., Lou, Y., & Yin, Y. (2011). *Proactive and Robust Dynamic Pricing Strategies for High-Occupancy-Toll (HOT) Lanes.* Paper presented at the Transportation Research Board 90th Annual Meeting.

Morgul, E. F. (2010). *Simulation based evaluation of dynamic congestion pricing algorithms and strategies.* Rutgers University Graduate School, New Brunswick.

Nagae, T., & Akamatsu, T. (2006). Dynamic revenue management of a toll road project under transportation demand uncertainty. *Networks and Spatial Economics, 6*(3-4), 345-357.

OCTA. (2003). *91 express lanes toll policy.* Orange County Transportation Authority Retrieved from http://www.octa.net//pdf/RevFinalTollPolicy7-30-03_v7.pdf.

Osorio, C., & Chong, L. (2015). A computationally efficient simulation-based optimization algorithm for large-scale urban transportation problems. *Transportation Science, 49*(3), 623-636.

Papageorgiou, M., Hadj-Salem, H., & Middelham, F. (1997). ALINEA local ramp metering: Summary of field results. *Transportation Research Record: Journal of the Transportation Research Board*(1603), 90-98.

Paz, A., Molano, V., & Gaviria, C. (2012). *Calibration of corsim models considering all model parameters simultaneously.* Paper presented at the 15th International IEEE Conference on Intelligent Transportation Systems.

Ramos, A. L., Ferreira, J. V., & Barceló, J. (2012). Modeling & simulation for intelligent transportation systems. *International Journal of Modeling and Optimization, 2*(3), 274.

Rathi, V., Antoniou, C., Wen, Y., Ben-Akiva, M. E., & Cusack, M. M. (2008). *Assessment of the Impact of Dynamic Prediction-Based Route Guidance Using Simulation-Based, Closed-Loop Framework.* Paper presented at the Transportation Research Board 87th Annual Meeting.

SANDAG. (2006). *Reversible lanes toll collection system documentation: request for proposals No. 5000680 for I-15 managed lanes toll collection system integration (design, build, operate, and maintain).* San Diego Association of Government, California
Retrieved from http://www.sandag.org/services/fastrak/pdfs/5000680/5000680_rfp.pdf.

Supernak, J., Kaschade, C., & Steffey, D. (2003). Dynamic value pricing on I-15 in San Diego: Impact on travel time and its reliability. *Transportation Research Record: Journal of the Transportation Research Board*(1839), 45-54.

Toledo, T., Mansour, O., & Haddad, J. (2015). Simulation-based Optimization of HOT Lane Tolls. *Transportation Research Procedia, 6*, 189-197.

Van Aerde, M., Hellinga, B., Yu, L., & Rakha, H. (1993). *Vehicle probes as real-time ATMS sources of dynamic OD and travel time data.* Paper presented at the Large Urban Systems-Proc. of the Advanced Traffic Management Conference.

Vaze, V., Antoniou, C., Wen, Y., & Ben-Akiva, M. (2009). Calibration of dynamic traffic assignment models with point-to-point traffic surveillance. *Transportation Research Record: Journal of the Transportation Research Board*(2090), 1-9.

Wang, Q. (1997). Using genetic algorithms to optimise model parameters. *Environmental Modelling & Software, 12*(1), 27-34.

Wang, Y., Zhang, G., & Northwest, T. (2009). A Self-Adaptive Toll Rate Algorithm for High Occupancy Toll (HOT) Lane Operations: Transportation Northwest, University of Washington.

Wen, Y., Balakrishna, R., Ben-Akiva, M., & Smith, S. (2006). *Online deployment of dynamic traffic assignment: architecture and run-time management.* Paper presented at the IEE Proceedings-Intelligent Transport Systems.

Wen, Y., Balakrishna, R., Ben-Akiva, M. E., & Smith, S. (2008). *Online Deployment of Dynamic Traffic Assignment: Evaluation and Lessons.* Paper presented at the Transportation Research Board 87th Annual Meeting.

Xu, S. (2009). *Development and test of dynamic congestion pricing model.* Master thesis, Massachusetts Institute of Technology.

Xu, Z., & Huang, A. Y. (2012). Safety benefits of converting HOV lanes to HOT lanes: Case study of the I-394 MnPASS. *Institute of Transportation Engineers. ITE Journal, 82*(2), 32.

Yang, L., Saigal, R., & Zhou, H. (2012). Distance-based dynamic pricing strategy for managed toll lanes. *Transportation Research Record: Journal of the Transportation Research Board*(2283), 90-99.

Yang, Q., Koutsopoulos, H., & Ben-Akiva, M. (2000). Simulation laboratory for evaluating dynamic traffic management systems. *Transportation Research Record: Journal of the Transportation Research Board*(1710), 122-130.

Yang, Q., & Koutsopoulos, H. N. (1996). A microscopic traffic simulator for evaluation of dynamic traffic management systems. *Transportation Research Part C: Emerging Technologies, 4*(3), 113-129.

Yildirim, M. B. (2001). *Congestion toll pricing models and methods for variable demand networks.* Phd thesis, University of Florida.

Yin, Y., & Lou, Y. (2009). Dynamic tolling strategies for managed lanes. *Journal of Transportation Engineering, 135*(2), 45-52.

Zhang, L., & Levinson, D. (2005). Road pricing with autonomous links. *Transportation Research Record: Journal of the Transportation Research Board*(1932), 147-155.

Zhou, X., & Mahmassani, H. S. (2006). Dynamic origin-destination demand estimation using automatic vehicle identification data. *Intelligent Transportation Systems, IEEE Transactions on, 7*(1), 105-114.

Zou, G., & Kulkarni, R. (2013). *A systematic genetic algorithm based framework to optimize intelligent transportation system (its) strategies.* Paper presented at the 20th ITS World Congress.