ABSTRACT

Title of Document:	AN INTEGRATED CONTROL MODEL FOR FREEWAY INTERCHANGES
	Zichuan Li, Doctor of Philosophy, 2011
Directed By:	Professor, Paul Schonfeld, Department of Civil & Environmental Engineering

This dissertation proposes an integrated control framework to deal with traffic congestion at freeway interchanges. In the neighborhood of freeway interchanges, there are six potential problems that could cause severe congestion, namely lane-blockage, link-blockage, green time starvation, on-ramp queue spillback to the upstream arterial, off-ramp queue spillback to the upstream freeway segments, and freeway mainline queue spillback to the upstream interchange. The congestion problem around freeway interchanges cannot be solved separately either on the freeways or on the arterials side. To eliminate this congestion, we should balance the delays of freeways and arterials and improve the overall system performance instead of individual subsystem performance.

This dissertation proposes an integrated framework which handles interchange congestion according to its severity level with different models. These models can generate effective control strategies to achieve near optimal system performance by balancing the freeway and arterial delays. The following key contributions were made in this dissertation:

1. Formulated the lane-blockage problem between the movements of an arterial intersection approach as an linear program with the proposed sub-cell concept, and proposed an arterial signal optimization model under oversaturated traffic conditions;

2. Formulated the traffic dynamics of a freeway segment with cell-transmission concept, while considering the exit queue effects on its neighboring through lane traffic with the proposed capacity model, which is able to take the lateral friction into account;

3. Developed an integrated control model for multiple freeway interchanges, which can capture the off-ramp spillback, freeway mainline spillback, and arterial lane and link blockage simultaneously;

4. Explored the effectiveness of different solution algorithms (GA, SA, and SA-GA) for the proposed integrated control models, and conducted a statistical goodness check for the proposed algorithms, which has demonstrated the advantages of the proposed model;

5. Conducted intensive numerical experiments for the proposed control models, and compared the performance of the optimized signal timings from the proposed models with those from Transyt-7F by CORSIM simulations. These comparisons have demonstrated the advantages of the proposed models, especially under oversaturated traffic conditions.

AN INTEGRATED CONTROL MODEL FOR FREEWAY INTERCHANGES

By

Zichuan Li

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2011

Advisory Committee:
Professor Paul Schonfeld, Chair
Professor Ali Haghani (ENCE, UMCP)
Professor Zhi-Long Chen (BMGT, UMCP)
Professor David J. Lovell (ENCE, UMCP)
Dr. Xianding Tao (ITS Division of the District Department of Transportation)

Acknowledgements

I would like to express my deepest gratitude to my advisor, Dr. Paul Schonfeld, for his inspiration, guidance, and encouragement during my study at the University of Maryland, College Park. Without his unconditional support and valuable advice, I would not be able to complete this dissertation. I may not be able to continue my study in the hardest time in my life without his support. Not only he inspires me with my research, but also he sets up a model for his student. His commitment to research, kindness to people, and unconditional support to students will inspire me for the rest of my life.

My special thanks go to the members of my doctoral examination committee, Dr. Ali Haghani, Dr. Zhi-Long Chen, Dr. David J. Lovell, and Dr. Xianding Tao, for their interests in my research, valuable suggestions, constructive comments and criticisms on this dissertation, which make this dissertation more solid and sound. I would thank Dr. Ali Haghani for his valuable support and understanding to a student in a hard time.

I would like to thank the State Highway Administration, Maryland for sponsoring part of this study under Project number MD-08-SP808B4D.

Finally, I would like to thank my family for their encouragement and understanding during these years. No words can express my appreciation to my beloved wife, Yu Lei, who has stayed with me for the last five years and sacrificed so much. I thank my brother, Han Li, who have assumed my responsibility to the family for the last five years. I appreciate my parents for their enormous love. I would like to thank my lovely son who brings me so much happiness during the preparation of this dissertation.

Table of Contents

List of Tables

List of Figures

Chapter-1:   Introduction

1.1. *Background*

The gridlock of urban transportation networks during peak hours has been frequently reported (Stringer 2006), and significantly disrupted urban transportation systems. Gridlock usually starts around an interchange or intersection. Signalized interchanges and intersections serve a critical function in urban transportation system. There are six potential problems arising around an oversaturated interchange, which are link blockage, lane blockage, green starvation, off-ramp spillback, on-ramp spillback, and freeway mainline spillback. Among them, the first three occur on the arterial, while the last three occur on the freeway. The following sections will discuss them in detail.

1.1.1 Arterial problems
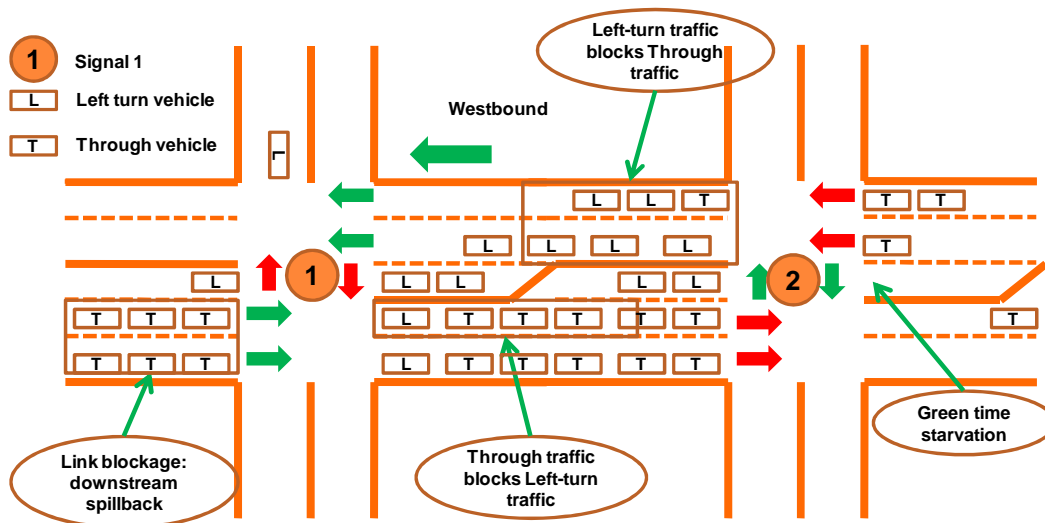
Figure 1-1 The blockage and starvation at oversaturated intersections

At many congested intersections, there exist the following two different blockage patterns (link spillback blockage and lane or movement blockage) and green starvation problems which are illustrated in Figure 1-1. Each is discussed in more detail below.

Link blockage

Link spillback blockage occurs when the queue from the downstream intersection spills back and thus blocks the upstream link traffic (see Figure 1-1). When link blockage occurs, the traffic on the upstream signal cannot move even during green phases. This is also called *De Facto Red* in the literature. In a busy urban transportation network, traffic could spill back further to several intersections upstream.

Lane blockage

Lane blockage (or movement blockage) may exist between different movements in the same approach if the queue exceeds its associated pocket. Figure 1-1 illustrates two types of lane blockage (left-turn traffic blocking the through flows, and the through traffic blocking the left-turn vehicles). Lane blockage is usually caused by improper signal timing or limited pocket space.

Green starvation

The green time starvation refers to the phenomenon in which a traffic stream is assigned more than enough green time, as illustrated by the westbound left-turn traffic of Signal 2 in Figure 1-1. Green starvation could and should be avoided by properly adjusting the signal timing.

1.1.2 Freeway problems

The arterial congestion could propagate to a freeway system during congested peak hours by blocking the connected off-ramp, which is termed off-ramp spillback. If the off-ramp queue further develops and blocks all freeway lanes, it becomes mainline spillback as illustrated in Figure 1-2. The congested freeway traffic could block its on-ramp traffic, which may further spillback to its upstream arterial and thus propagate congestion to the associated arterial. Those phenomena have been observed by us on the Capital Beltway in Washington, D.C.
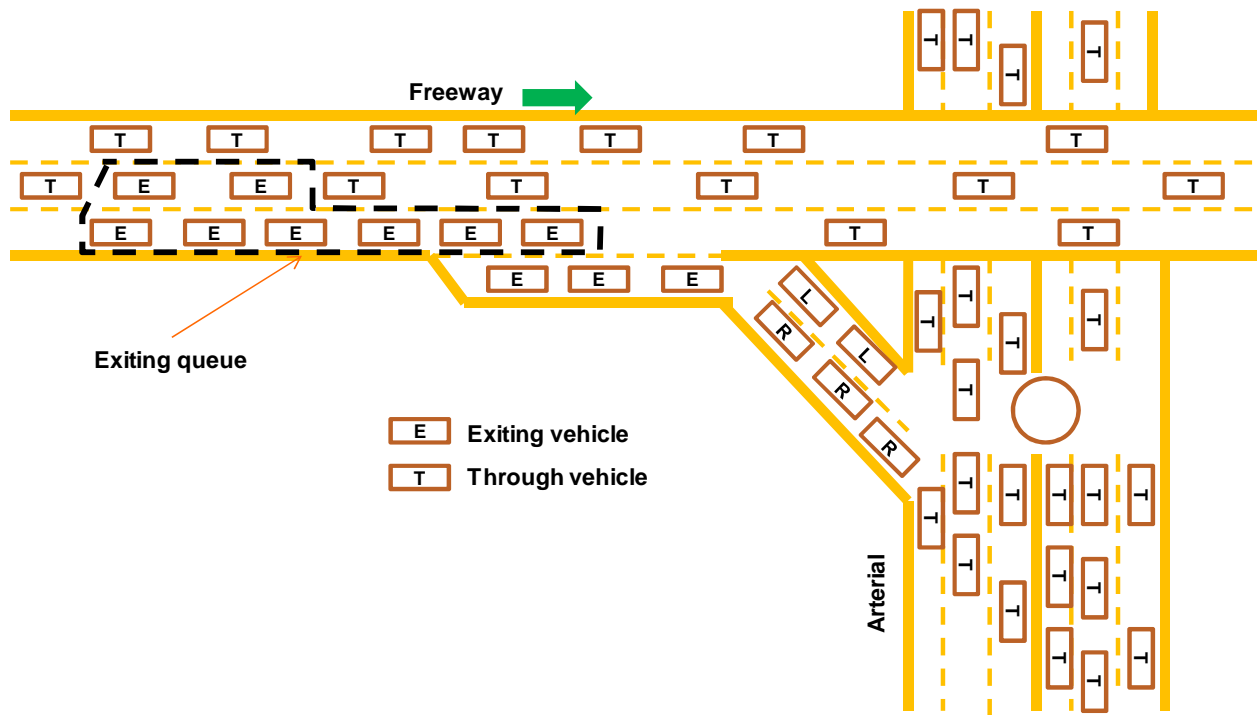
Figure 1-2 Problems for an oversaturated signalized interchange

Off-ramp spillback

When the exit volume from the freeway increases, the exit queue eventually spills backs to its immediate freeway mainline. Off-ramp spillback diminishes freeway through capacity and thus causes freeway congestion. Off-ramp spillback has also been reported by researchers (Cassidy et al. 2002; Jia et al. 2004; Lovell 1997).

3

On-ramp spillback

The on-ramp metering system normally benefits freeway mainline traffic by controlling the volume entering from an on-ramp (Papageorgiou and Kotsialos 2002). However, if the on-ramp demand exceeds the metering rate, the on ramp queue could block the upstream arterial and diminish its through capacity, thereby increasing arterial delay.

Freeway mainline spillback

In some extreme scenarios, the off-ramp queue could develop further and block all freeway lanes. Freeway mainline queue will form if there is no sufficient through capacity. The freeway mainline queue could block the upstream on-ramp traffic and propagate congestion to the upstream interchange.

To sum up, all the aforementioned problems could cause gridlock in a transportation network. Those problems could occur separately or concurrently in some scenarios. In current practice, the freeway and arterial systems are controlled and managed separately. Under that separate operation framework, the off-ramp and on-ramp spillback problems are difficult to handle. To address those issues and improve overall system performance, it is essential to analyze a freeway and its adjacent arterial in an integrated way.

However, there are limited studies (Cremer and Schoof 1989; Papageorgiou 1995; Tian et al. 2002; van den Berg et al. 2003; Wu and Chang 1999) which attempt to integrate control both freeways and arterials. Hence, in the course of developing an effective integrated control strategy, many critical theoretical and operational issues should be explored. Some of those issues include:

➤ How to model the link- and lane-blockage for an oversaturated arterial intersection;

➢ How to model freeway traffic, capture on-ramp, off-ramp, and freeway mainline

spillback under oversaturated condition, and naturally downgrade to normal status for

undersaturated conditions;

➢ How to choose a proper control model to balance the computation complexity and system

requirements;

➢ How to balance the delay between freeways and arterials to achieve a system optimal

solution;

➢ How to solve the proposed model efficiently and sufficiently quickly, and;

➢ How to achieve robust control even in the presence of traffic measurement errors and

disturbances.

1.2. *Research objectives*

The primary objective of this dissertation is to develop an integrated traffic control

system for freeways and arterials, which can assist traffic operation practitioners to increase the

traffic system's performance and reduce travel delay. Specifically, the proposed system should

be able to:

➢ Generate optimal or near optimal arterial signal timings for arterial intersections under

both oversaturated and under-saturated traffic conditions;

➢ Produce integrated control strategies for arterial signals and interchange signals with or

without presence of off-ramp spillback, on-ramp spillback, or freeway mainline spillback;

➢ Balance freeway and arterial delays to achieve a system optimal for oversaturated

conditions.

To accomplish the above objectives, this proposed system should be able to:

➢ Capture the complex arterial traffic interactions under oversaturated conditions, i.e., account for link and lane blockage in the same modeling framework;

➢ Represent the traffic evolution around an oversaturated interchange under off-ramp spillback, on-ramp spillback, or freeway mainline spillback circumstances;

➢ Generate optimized solutions efficiently and robustly for a realistic size network, and;

➢ Switch between different control models smoothly and automatically based on the traffic pattern.

1.3. *Dissertation organization*

Based on the proposed research objectives, this dissertation has organized the primary research tasks into nine chapters. Those chapters are outlined and their relations are illustrated in Figure 1-3.

The remaining chapters of this dissertation are organized as follows:

➢ Chapter 2 presents a comprehensive literature review of existing studies on various control models for freeways and arterials. This literature review aims to explore the core concepts of existing models and identify their advantages and limitations, their evolution, and their potential enhancements.

➢ Chapter 3 provides the framework of the proposed integrated control system. This chapter presents the key research issues first based on the aforementioned system features followed by a system operation flowchart, lists the system key inputs, and describes the functions of each principal module in detail. After that, it presents the key modeling issues. Finally, the last section concludes this chapter.

```
┌─────────────────────────────┐        ┌─────────────────────────────┐
│ Chapter 1: Introduction     │        │ Chapter 3: System framework │
└─────────────────────────────┘        │ • Identify research issues  │
              │                        │ • System operation flowchart│
              ▼                        │ • Principal module functions│
┌─────────────────────────────┐        └─────────────────────────────┘
│ Chapter 2: Literature review│                      │
└─────────────────────────────┘                      │
```

┌──────────────────────────────────────┐   ┌──────────────────────────────────────┐
│ Chapter 4: Arterial signal   │   │ Chapter 6: Numerical case studies   │
│ optimization model   │   │ • Arterial signal optimization model │
│ • Lane-blockage model   │   │ • Integrated interchange control model│
│ • Arterial signal optimization model │   └──────────────────────────────────────┘
│ • Solution algorithm   │
└──────────────────────────────────────┘   ┌──────────────────────────────────────┐
                                             │ Chapter 7: Multi-interchange integrated│
┌──────────────────────────────────────┐   │ control model   │
│ Chapter 5: Integrated interchange   │   │ • Freeway through model with queue   │
│ control model   │   │   presenting   │
│ • On-ramp spillback model   │   │ • Integrated multi-interchange control│
│ • Off-ramp spillback model   │   │   model   │
│ • Integrated single interchange model │   └──────────────────────────────────────┘
└──────────────────────────────────────┘
                                             ┌──────────────────────────────────────┐
                                             │ Chapter 8: Hybrid genetic algorithms │
                                             │ • GA   │
                                             │ • Simulated annealing-GA   │
                                             └──────────────────────────────────────┘

                          ┌──┐
                          │ Chapter 9: Conclusions and recommendations│
                          └──┘

Figure 1-3 Dissertation Organization

➢ Chapter 5 proposes an integrated single-interchange control model, which optimizes the interchange signals and their immediate upstream and downstream signals jointly with the presence of on-ramp and off-ramp queue spillback. The proposed model analyzes arterial links with the model proposed in Chapter 4, and represents the on-ramp and off-ramp queue spillback dynamic with a set of new formulations. The proposed model is solved with a similar GA-based algorithm.

➢ Chapter 6 presents sensitivity analysis results for models in Chapter 4 and Chapter 5. This chapter compares model performance with an independent microscopic simulation package, TSIS/CORSIM. The experiments include the comparison between the proposed arterial signal optimization model and Transyt-7F, the integrated control model for single interchange and Transyt-7F model, and the arterial signal optimization model and the integrated interchange control model.

➢ Chapter 7 deals with the freeway mainline spillback problem. This chapter proposes a freeway queue spillback module under the Cell Transmission framework, which enables the multi-interchange integrated control model to represent the freeway mainline queue spillback dynamics. The proposed model can be solved with a GA-based algorithm.

➢ Chapter 8 attempts to improve the efficiency of the GA algorithm by integrating it with two other optimization algorithms: simulated annealing (SA) and simultaneous perturbation stochastic approximation (SPSA). This chapter focuses on developing two algorithms which are: pure SA and SA-GA, and comparing their performance in arterial signal optimization model. The arterial segment connected to the interchange of I-495/MD97 serves as the field site for numerical comparison.

➢ Chapter 9 summarizes the contributions of this dissertation and provides recommendations for future research.

Chapter-2: Literature Review

2.1. *Introduction*

This chapter summarizes major literature findings from recent decades on various aspects of traffic corridor management during recurrent traffic congestion. It aims to not only highlight the critical issues but also identify potential research directions for this research field. We start with interchange control models (Section 2.2), and then discuss integrated corridor control models (Section 2.3). As an important component of the integrated control model, the off-ramp queue will affect through traffic behavior when spillback. We discuss the existing studies considering lateral friction to drivers' behavior in Section 2.4. This Chapter also includes a discussion in the solution algorithms for signal optimization models in Section 2.5. To make this review more comprehensive, this chapter also reviews freeway ramp metering and arterial signal control models, which are the two major components of the freeways and arterials integrated control model.

2.2. *Interchange Control Models*

Interchange control refers to joint control of the interchange signals and on-ramp meters in the same model to maximize the system performance. The literature has some research on interchange operations (Chlewicki 2003; Dorothy et al. 1998; Messer and Berry 1975; Messer et al. 1977; Munjal 1971; Radwan and Hatton 1990). PASSER III (Venglar et al. 1998) is a computer model specifically designed to analyze the operation of an isolated interchange. However, these models are limited to under-saturated conditions, and do not consider off-ramp queue spillback, on-ramp queue spillback, or the link queue blockage problem.

For congested interchanges, Kovvali et al. (2002) introduce a GA-based model to optimize the diamond interchange signal timing for both under-saturated and oversaturated traffic conditions while considering link queue spillback. Engelbrecht and Barnes (2003) evaluate eight controller features, which can potentially improve diamond interchange operations under certain conditions, by CORSIM simulation combined with hardware-in-the-loop, and find that the separate intersection diamond control model has the potential to provide more efficient control than the three-phase or four-phase sequences that are typically used for diamond interchange control in Texas. Tian et al. (2004) consider the on-ramp spillback problem by integrating on-ramp metering and interchange signals optimization. Fang and Elefteriadou (2006) propose an adaptive algorithm based on a forward dynamic programming method, which yields promising results for fluctuating demand for congested conditions. Lee et al. (2006) investigate the performance of the three-phase and four-phase with two-overlap phasing strategies for diamond interchanges under congested conditions using CORSIM combined with hardware-in-the-loop technology, and conclude that the performance of each phasing strategy for congested traffic conditions is dependent on the traffic pattern and ramp spacing. Li et al. (2009) proposed a model to prevent off-ramp queue spillback using the traditional Cell-Transmission by controlling the adjacent arterial signals. The CORSIM simulation results indicate that this model is promising. Zhang et al. (2009) propose a local synchronization control scheme to manage queues at critical locations based on the traditional Cell-Transmission. This model attempts to distribute queues over a wider area through coordination, which involves on-ramp priority control, off-ramp priority control, and internal metering.

Despite all the progress on interchange signal optimization, no existing model considers the aforementioned issues jointly, especially the off-ramp queue spillback effect. Those models intend to improve the freeway performance rather than that of the overall system.

2.3. *Integrated Corridor Control Models*

Integrated corridor control models control freeway and its parallel arterials jointly in order to improve system-wide performance. These models optimize the off-ramp diversion rates, on-ramp metering rates, and arterial signal timings to achieve optimal system performance. Early studies in this category focus mainly on modeling and simulation analyses (Reiss 1981; Van Aerde and Yagar 1988). Few analytical studies attempt to deal with integrated controls for freeways and arterials, which are summarized below.

Cremer and Schoof (1989) first formulate an integrated control model to jointly control four types of traffic controls, including off-ramp traffic diversion, on-ramp metering, mainline speed limit, and arterial signal. This model represents freeway traffic by continuous flow model and arterial traffic by the platoon dispersion model of Transyt. A mixed integer nonlinear optimal control model is formulated to minimize the total delay time for the entire corridor system, and solved with a heuristic decomposition method. Van Den Berg et al. (2004) propose a model predictive control (MPC) approach for freeways and arterials based on enhanced macroscopic traffic flow models. This study minimizes the total system travel time and solved with a predictive control framework. The proposed model represents freeway dynamics with METANET (a continuous flow model), and arterial dynamics, on-ramps and off-ramps with a enhanced Kashani model (Kashani and Saridis 1983).

11

A dynamic system-optimal control model (Chang et al. 1993) minimizes the total travel time for a commuting corridor, including a freeway and its parallel arterials. However, this model assumes that travel times and queue lengths are known, which is unrealistic. Papageorgiou (1995) develops a linear optimal control based on the store-and-forward model, which can be applied to both motorways and signal-controlled urban roads with some approximation. Wu and Chang (1999) present a set of linear programming (LP) models to optimize control strategies for commuting corridors under non-recurrent congestion situations, which is solved by a heuristic solution algorithm duo the number of LPs may be very large.

To sum up, the integrated corridor control models jointly optimize the arterial signals, off-ramp diversion rate, and on-ramp metering rate under various traffic conditions in terms of overall system performance. Compared with interchange control, these models have a wider control boundary and usually include several interchanges and one or more parallel arterials. The formulations turn out to be large-scale and nonlinear, and thus very difficult to solve with traditional algorithms. However, these models emphasize the performance balance of freeways and arterial without considering lane-blockage, off-ramp spillback, and on-ramp spillback problems, which makes their results unreliable under oversaturated conditions.

2.4. *Car-following models with lateral friction*

On a congested freeway, the behavior of drivers traveling with a long exit queue in the neighboring lane can be affected by the exit queue in the following two aspects. Firstly, the lane-changing from the exit queue to the through lanes makes the drivers in the moving lane more cautious, thus reducing the moving lane traffic speed; secondly, drivers feel unsafe driving at high speed with a long vehicle queue due to some complex psychological mechanism. Those two

factors could reduce the capacity of the through lanes. However, there is very little useful literature on this topic.

One of the key assumptions of traditional car following models is that vehicles are influenced only directly by the leading vehicle, which means no lateral friction is considered. These models assume ideal conditions, including each vehicle positioned in the center of the lane, sufficient lane width, clear road markings, and good visibility. However, the drive environment is far from ideal in reality. The driver behavior of a following car can be affected by other factors which have been reported in the literature.

Case et al. (1953) find that drivers intend to avoid objects near their path located on the roadside, and change their behavior for different curb height, should width, and lane width. Taragin (1955) reports that drivers intentionally keep away from traffic on neighboring lanes as volume increases. On a four-lane highway, vehicles in the shoulder lane travel closer to the shoulder, and those in the median lane travel closer to the median on average. May (1959) introduces four types of friction in traffic flow, as internal, medial, marginal, and intersectional. May also defines interval friction as the friction between vehicles moving in the same direction, which can be influenced by the number and width of lanes, horizontal and vertical alignment, and uniformity and smoothness. However, May provides no mathematical formulation for this effect. Michales and Cozan (1963) investigate the speed and lateral clearance relationship of a vehicle, and report that a channel constructed with plastic cones can be used to control traffic speed. The factors affecting the speed of traffic include the channel width and the longitudinal distances of between consecutive cones. Gunay (2003) employs the terminology lane-based-driving discipline to represent the tendency to drive within a lane by keeping to the centre as closely as possible, and proposes five possible methods to quantify it.

Despite all the efforts in this respect, there is no existing mathematical model which can consider the effect on driving behavior of the lateral friction, especially the effect of an exit queue on drivers in the adjacent freeways lanes. Without considering the lateral friction, the through traffic capacity with off-ramp queue spillback could be overestimated, thus compromising the credibility of the final optimized results.

2.5. *Solution Algorithms for Signal Optimization Problem*

To improve signal timings for oversaturated intersections, considering multiple adjacent intersections simultaneously is usually required due to link spillback from downstream intersections. Models in this group are usually called system-wide control models (Spall and Chin 1997). They adjust all signal timings of a network to improve overall system performance in response to instantaneous changes of traffic conditions. Different heuristic algorithms have been adapted to solve those models. Among those algorithms, GAs are among the most popular.

The basic GA has been applied to various models (Ceylan and Bell 2004a; Ma and Abdulhai 2002; Rahmani et al. 2011). For a signal optimization problem, Hadi and Wallace (1993) combine GA and hill-climbing algorithm of Transyt-7F to optimize all the signal design elements. In their model, the main purpose of the GA is to optimize the signal sequence. They propose two alternatives. The first executes GA and hill-climbing concurrently, and uses GA to optimize phase sequences and cycle lengths. The other runs the GA and hill-climbing sequentially. The GA first optimizes the cycle length, phase sequence, and offset, and then hill-climbing is used to adjust the resulting signal timing. A fraction-based signal decoding scheme has been suggested by Park et al (1999) and employed in various signal optimization models (Li

2011; Park et al. 2000; Stevanovic et al. 2007; Stevanovic et al. 2008). Other applications of

GA's include the signal optimization models of Lo et al. (2001; 2004).

Simulated Annealing (SA) is another popular heuristic optimization algorithm, which

was introduced by Kirkpatrik et al. (1984). SA belongs to a class of algorithms named

probabilistic hill-climbing algorithms. It is a general purpose optimization method reflecting an

analogy between the solutions of an optimization problem and the energy states of a slowly-

cooled physical solid. SA is capable of obtaining solutions near global optima. These solutions

do not depend on the initial solutions as do other algorithms such as hill climbing. Hadi and

Wallace (1994) first applied the fast simulated annealing (Szu and Hartley 1987) algorithm to

signal optimization.

While GA's are very popular for oversaturated signal optimization, they suffer from the

following problems. One problem frequently found in GA optimization is premature

convergence, which is typically the result of the extreme reliance on crossover. The dominance

of crossover can result in stagnation as the population becomes more homogeneous and the

mutation rate is too low to search other solution domain. A second well-known problem of GA's

is their poor hill-climbing performance. After finding a near-optimal solution, a GA often has

difficulty converging to the optimal solution. A third complaint about GA's is their large

memory requirement. Since a GA should maintain a large population of solutions, it uses much

memory as the problem dimensions increase.

2.6. *Freeway Traffic Access Control Strategies*

The freeway control strategies in the literature can be classified into two major categories:

access control and link control. The access control includes various on-ramp control strategies,

which limit on-ramp volume to keep the downstream freeway volume under its capacity. The link control affects freeway drivers' behavior by suggesting or enforcing variable travel speeds through Variable Message Signs (VMS), or providing travel time information etc. This study mainly deals with access control, so the following section will focus on on-ramp metering strategies.

The basic idea of most on-ramp metering methods is to maintain free flow on the freeway mainline bottleneck (Papageorgiou and Kotsialos 2002) by limiting on-ramp volume, thus increasing freeway mainline speed and improving safety. The on-ramp metering strategies generally result in a more efficient use of the existing freeway infrastructure and benefit mainline traffic(Arnold 1998). The existing ramp metering methods can be classified into three categories: pre-timed metering, which derives ramp metering rates based on the historical average volumes without considering real-time measurements; automatic metering, which keeps freeway traffic measurements close to their pre-specified values by adjusting ramp metering rates; and adaptive ramp metering, which determines the ramp metering rates based on pre-set objective function(s).

2.6.1 Pre-timed Metering Strategies

Pre-timed metering strategies are the simplest on-ramp metering strategies, which usually provide different timings based on the average demand pattern. Wattleworth (1963) is one of the pioneering studies in this category. It considers the physical upper- and low-bounds of each ramp's metering rate, and maximizes the total entry flow under the constraints of freeway mainline capacity. Similar studies (Chen et al. 1974a; Chen et al. 1974b; Schwartz and Tan 1977; Tabac 1972; Wang 1972; Wang and May 1973; Yuan and Kreer) formulate linear or quadratic programming problems with different objective functions to optimize ramp metering rates.

Papageorgiou (1980) proposes a Linear Programming (LP) on-ramp metering optimization model, which considers the travel time from upstream to the downstream by assuming a constant travel time. The proposed model is solved with a decomposition approach. Lovell and Daganzo (2000) improve time-dependent control strategies for small freeway networks with bottlenecks and unique origin-destination path, which requires time-dependent bottleneck capacities and is solved with a greedy heuristic algorithm.

Pre-timed ramp metering strategies optimize ramp metering rates under average traffic pattern. However, the actual traffic demand varies from average pattern frequently. Therefore, pre-timed ramp metering strategies may lead to non-optimal results, such as underutilizing the freeway mainline capacity or allowing mainline queue formation, even though they provide optimal on-ramp metering rate for average traffic patterns.

2.6.2 Automatic metering strategies

With the help of real-time traffic sensor data, the automatic metering strategies optimize on-ramp metering rates according to actual traffic conditions. Existing automatic ramp metering strategies can be classified into the following two categories: local (or isolated) on-ramp metering models, which determine the ramp metering rate based solely on the adjacent freeway mainline traffic conditions; and coordinated on-ramp metering models, which decide the metering rates with upstream and downstream on-ramp conditions.

Local automatic on-ramp metering strategy

The local ramp-metering strategy determines the metering rate solely based on the local traffic information. Existing models of this type can be further classified into volume-based models and occupancy-based models.

Volume-based models

Volume-based strategies compute on-ramp metering rate based on upstream freeway volume and downstream through capacity. The models in this category include the demand-capacity model (Masher et al. 1975), zone control model (Stephanedes 1994; Xin et al. 2004), and congested pattern control model (ANCONA)(Kerner 2005).

The *demand-capacity model* is proposed by Masher et al. (1975). This model attempts to fully utilize the downstream freeway mainline capacity (q_{cap}), which sends maximum on-ramp rate provided that the downstream freeway mainline volume is under its capacity. The demand-capacity model reduces the on-ramp metering rate to the minimum, r_{min}, to avoid congestion if the downstream occupancy exceeds its preset threshold. This model is not a closed-loop but an open-loop disturbance-rejection model which is quite sensitive to disturbances. The *zone control model* (Stephanedes 1994) is another volume-based model which has been employed by Minnesota Department of Transportation for many years. It divides the mainline freeway into several zones and each zone has no more than one on-ramp. This model is further extended to the stratified ramp control model of Xin et al.(2004) , which considers the on-ramp demands and queue lengths. The *congested pattern control model (ANCONA)* proposed by Kerner (2005) keeps on-ramp bottleneck at the minimum possible level without propagating upstream. This method is based on three-phase traffic theory.

Occupancy-based strategy

Occupancy-based strategy determines the ramp metering rate based on the occupancy of downstream freeway mainline and uses feedback regulation to maintain a pre-specified occupancy. ALINEA (Papageorgiou et al. 1991), Neural control algorithm (Xin et al. 2004;

Zhang and Ritchie 1997), and Iterative Learning Approach (Hou et al. 2008) are examples in this category.

The ALINEA algorithm (Papageorgiou et al. 1991) is a closed-loop ramp metering strategy using the classical feedback theory. It adjusts the metering rates in response to even slight differences of the target occupancy and measured occupancy. It can maintain high volume in the freeway mainline without causing congestion. The local artificial neural network method (Zhang and Ritchie 1997) is derived from the fundamental diagram, which models the ramp metering problem as a nonlinear feedback control problem with artificial neural networks, which are composed of one or several multilayer feed-forward neural networks. The iterative learning approach (Hou et al. 2008) formulates the density-based ramp metering control problem as an output tracking and disturbance rejection problem, and employs the iterative learning control combined with error feedback to achieve system robustness.

Both the ALINEA and Neural control algorithms are effective, robust, and flexible for moderate congestion. They are easy to implement since the only parameters are the target occupancy. However, neither considers on-ramp queue spillback directly, and both would have difficulty in balancing freeway and on-ramp congestion.

Coordinated automatic metering strategies

The coordinated metering strategies determine the metering rate for each on-ramp based on not only adjacent traffic conditions but also system-wide traffic conditions. Based on the way to consider local and system-wide traffic conditions, the coordinated metering strategies can be divided into three groups: *cooperative ramp-metering*, which adjusts the resulted metering rate from the local traffic condition according to the system-wide conditions; *competitive ramp-*

metering, which first computes two sets of metering rates based on local and system-wide traffic condition separately, and then chooses the lower one one as the final results; *integral ramp-metering*, which determines all the ramp metering rates by system-wide traffic conditions simultaneously. Extensive summaries of those strategies are available in the literature (Bogenberger and May 1999; Jacobson et al. 1989a; Nihan and Berg 1991; Zhang et al. 2001). Although these strategies are limited in many aspects, some field experiments and many simulation studies have confirmed that they can reduce delay successfully (Bogenberger and May 1999).

Cooperative ramp-metering

The cooperative ramp metering algorithms first determine on-ramp metering rate based on local traffic conditions, and then adjust the rate based on system-wide information to avoid both freeway mainline congestion and on-ramp queue spillback. The helper ramp metering algorithm (Lipp et al. 1991) and linked-ramp algorithm (Banks 1993) are examples in this group.

The *helper ramp metering algorithm* (Lipp et al. 1991) consists of a local traffic responsive metering algorithm and a system-wide coordinated control. The local responsive algorithm selects one of six pre-defined metering rates based on its upstream mainline occupancy for each on-ramp. The coordinated control part of the algorithm will override upstream ramp meter if a meter reaches critical status. The *Linked-ramp algorithm* (Banks 1993) is based on the demand-capacity concept. This algorithm first set the on-ramp metering rate to the difference of the pre-specified target flow rate and its upstream freeway flow rate. When the metering rate of a ramp is among the lowest three rates pre-specified by the control algorithm, the coordinated control strategy of linked-ramp algorithm reduces its next upstream ramp metering rate to the same, and propagates this rate reduction procedure upstream.

Competitive ramp-metering

The competitive algorithm computes two sets of metering rates based on local and system-wide conditions respectively, and applies the more restrictive one. The example algorithms include the bottleneck algorithm (Jacobson et al. 1989b) and system-wide adaptive ramp metering (SWARM) (Ahn et al. 2007; Paesani et al. 1997).

The bottleneck algorithm uses upstream occupancy data and bottleneck data to determine both a local and a bottleneck metering rate, and then select the more restrictive rate. The local rate is set to be the difference between the measured upstream volume and the capacity estimated from historical data for each ramp. The bottleneck algorithm identifies bottlenecks based on historical traffic conditions, determines its volume reduction based on flow conservation, and then distributes the volume reduction to upstream ramps according to pre-specified weights.

The SWARM algorithm operates at two levels. The local level determines metering rate based on local density. The system-wide control decides the overall reduction for the upstream ramps of a critical bottleneck, and then applies pre-determined weights to distribute the volume to upstream ramps, which can be employed to obtain a new set of ramp metering rates. In the SWARM algorithm the bottlenecks are identified based on predicted traffic conditions rather than measured traffic conditions.

Integral ramp-metering
Integral ramp metering algorithms directly generate the metering rates from the system-wide information. The METALINE, fuzzy logic algorithm, and coordinated artificial neural network algorithm belong to this category.

METALINE (Papageorgiou 1990) is usually viewed as a generalization and extension of ALINEA, which is theoretically sound, robust, and easy to implement. However, it is a challenge to find the proper control parameters matrices and the target occupancy vector. Fuzzy logic algorithms (Chen et al. 1990; Meldrum and Taylor 1995; Sasaki and Akiyama 1986; Taylor et al. 1998) convert empirical knowledge about traffic flow ramp control into fuzzy rules. However, their system-wide control rules may be quite complex and are not always straightforward. It often is a challenge to calibrate the parameters which usually depend on traffic conditions. The coordinated artificial neural network algorithm (Wei and Wu 1996) divides the controlled freeway into control zones, and represents its system-wide effects to a particular on-ramp by a "hidden layer".

2.6.3 Adaptive ramp-metering strategy

Adaptive ramp metering algorithms have an explicit objective function linked to the control strategy. These objective functions include minimizing total travel time and maximizing system throughput. Some of those algorithms are presented as follows.

Linear programming algorithms are among the earliest adaptive ramp-metering methods for both practice and research. The Hanshin algorithm (Yoshino et al. 1995) employs a linear programming model to maximize the total number of vehicles entering the system while preventing traffic congestion in any section of the freeway mainline and avoiding negative effects on the adjacent arterials. Recognizing the complexity of formulation and the difficulty in obtaining real-time OD information, Zhang and Levinson (2004) formulated the optimal ramp control problem as a linear programming whose input variables are all directly measurable by detectors in real time. Gomes and Horowitz (2006) formulates ramp metering problem as a linear

programming, which minimizes the total travel time and models the traffic dynamic by the asymmetric cell transmission model (ACTM).

The Dynamic metering control algorithm (Chen et al. 1997) consists of four operational elements: state estimation, OD prediction, local metering control, and area-wide metering control. The basis of this algorithm is formed by the hierarchical structure including a local feedback control algorithm (ALINEA) and a system-wide control model. The system-wide control model employs a linear-quadratic feedback control model to produce nominal metering rate for the local controllers, which then compensate the nominal set values based on local traffic disturbances and prediction errors.

The Linear-quadratic (LQ) feedback control algorithm is one of the most commonly studied methods within automatic control theory for coordinated ramp metering (Kaya 1972; Papageorgiou 1983; Papageorgiou et al. 1990; Payne et al. 1985; Yuan and Kreer 1971). The LQ feedback strategy linearizes the nonlinear model equations around a certain desirable trajectory and employs a quadratic penalty function in the objective function to represent the state and control deviations from the desired trajectory.

The Rolling time horizon or successive optimization algorithm is another way to solve the large scale ramp metering optimization problem with real-time detector information. Chang et al. (1994) present an algorithm to capture the dynamic evolution of traffic with a two-segment linear flow-density model, and employ a successive linear programming algorithm to determine optimal metering rates. The model has been integrated with INTRA, a microscopic freeway simulation model, for simulation experiments under non-recurrent congestion conditions.

In summary, ramp metering is one of the most direct and efficient ways of mitigating freeway congestion if appropriately implemented. The benefit includes increased freeway mainline throughput and decreased total travel time. However, this may be achieved at the cost of excessive queues at the on-ramp, which may spill back and block neighboring urban arterials (Levinson and Zhang 2006). To achieve a better performance for the overall transportation network, the control boundary should be extended to cover both freeways and arterials.

2.7. *Arterial Traffic Signal Control Strategies*

Signal control has been widely implemented as an effective strategy to increase arterial capacity and mitigate congestion. The earliest work in signal control is attributed to Webster and Cobbe (1967), who introduce a formula for optimizing the signal timings for an isolated intersection.

The existing signal optimization models in the literature can be classified into two categories: undersaturated models and oversaturated models. Despite the large body of literature related to signal control, most existing studies focus on undersaturated traffic conditions. The following section will review key models for arterial signal optimization for undersaturated and oversaturated traffic conditions separately.

2.7.1 Undersaturated signal control models

The undersaturated condition of an intersection refers to those traffic conditions under which there is no cycle failure where no queues remain at the ends of green phases. Most available signal control models in the literature are for undersaturated conditions.

Phase-base models

The phase-base models are among the earliest version of signal optimization models, which are pre-timed (or fixed-time) models for isolated intersections. Models in this category optimize the splits, cycle length, and phasing to minimize the total delay or maximize the intersection capacity. SIGSET (Allsop 1971) and SIGCAP(Allsop 1976) are well-known examples in this category. SIGSET employs the nonlinear total delay function by Webster (1958) for undersaturated conditions to calculate intersection delay. SIGSET considers constraints including discharge capacity for each phase, maximum cycle length and minimum green time. The resulting model is a nonlinear programming with linear constraints. SIGCAP maximizes the multiply of the demand pattern under identical constraints of SIGSET. This concept of maximizing the multiply is extended to the reserve capacity model (Wong et al. 2007; Wong and Yang 1997). A binary-mixed-integer version of the phase-based model which considers different staging combinations is proposed by Improta and Cantarella (1984). The phase-based strategies are only suitable for undersaturated isolated intersection signal optimization.

Bandwidth models

The bandwidth refers to the portion of a signal cycle during which a car (traveling at pre-assigned speeds) could start at one end of a street and can reach the other end without stopping for a red light (Morgan and Little 1964). A mixed-integer linear programming proposed by Little et al. (1981; 1966) is the first version of MAXBAND. In MAXBAND, green splits are considered as input. Therefore, the problem transforms into finding the offsets of the arterial signals to maximize the inbound and outbound bandwidths. This mixed-integer linear programming is solved by the traditional branch-and-bound algorithm and a more efficient solution method is proposed by Chaudhary et al. (1991). This model is further extended to

25

optimize left-turn phase sequence (Chang et al. 1988), to optimize the weighted bandwidth for each directional road section(Gartner et al. 1991), to consider time-varying demand (Han 1996), and to optimize arterial networks (Stamatiadis and Gartner 1996).

Platoon dispersion model (Transyt)

Transyt (Robertson 1969) is the most known and applied signal control software package, and it is often a reference signal control strategy for testing improvements enabled by real-time strategies (Papageorgiou et al. 2003). Transyt employs the concept of "platoon dispersion" to model flow propagation along a link. The core concept of platoon dispersion model can be expressed as follows (Seddon 1972):

$$q_t' = \sum_{i=t}^{\infty} F(1-F)^{i-T} q_{t-1} \tag{2-1}$$

$$F = \frac{1}{1 + \alpha\beta T_a} \tag{2-2}$$

where t is time interval index; q_t' is the arriving flow at the downstream of the link at time t; q_{t-1} is the discharging flow at upstream of the link at time t-1; F is the smoothing factor; T is the minimum travel time for the link, measured in units of time steps; T_a is the mean roadway travel time measured in units of time steps; α is platoon dispersion factor, which is allowed to vary between 0.2 to 0.5 depending on the level of friction along the roadway; β is the travel time factor, which has been fixed as 0.8 by (Robertson 1969). Intensive research (Farzaneh and Rakha 2006; McCoy et al. 1983) has been done on calibrating α and β.

All aforementioned models are pre-timed signal optimization models, based mainly on historical average traffic conditions. The traffic-responsive version of Transyt, SCOOT(Split,

Cycle, Offset Optimization Technique), has been proposed by Hunt et al. (1982) and extended

later in several respects. Since such off-line models lack the functions to respond to traffic surge

or fluctuation in real time, some researchers have later worked on developing real-time adaptive

control systems. Example of such system include: SCAT(Sims and Dobinson 1980),

OPAC(Gartner 1983), PRODYN(Henry et al. 1984), CRONOS(Boillot et al. 1992),

RHODES(Sen and Head 1997), ARTC (Kim et al. 1993). Those adaptive system control systems

are beyond the scope of this study, and are not reviewed in this chapter.

2.7.2 Oversaturated signal control models

Oversaturation refers to those traffic conditions under which traffic queues persist from

cycle to cycle either due to insufficient green splits or blockage of adjacent traffic movement.

Under those conditions, queues along signalized arterials may block upstream intersections, thus

exacerbating the already bad conditions. Adaptive signal control strategies will not work very

well under oversaturated traffic conditions since they general just operate at maximal call mode

and it is more useful for highly varying traffic conditions(Lo et al. 2001). The earliest research

on oversaturated signal control models goes back to Gazis and Potts(1963), who optimized two

closely spaced and oversaturated intersections using a graphic methods. Recently, oversaturated

signal optimization problems have attracted increasing attention.

Queue polygon approach

The queue polygon approach analyzes the vehicle trajectory (or time-space diagram) of

the end-of-queue vehicle to compute the delays of the entire queue. The trajectories of individual

vehicles in motion are portrayed by sloping lines in a diagram that has a horizontal time axis and

a vertical axis representing the distance from a reference point.

Abu-Lebdeh and Benekohal (1997) develop an algorithm for optimizing the signal timing for oversaturated arterials with a queue polygon approach, which assumes continuous queue and considers the link blockage problem (De Facto Red) problem. The resulting model is solved with a GA-based algorithm. This model is further extended to include a disutility function, which enables it to evaluate a variety of traffic management scenarios (Abu-Lebdeh and Benekohal 2000; Abu-Lebdeh and Benekohal 2003). Abu-Lebdeh et al. (2007) propose models that capture traffic output of intersections under congested interrupted flow conditions with explicit consideration of interactions between traffic streams at successive signals.

Chang and Lin (2000) analyze the queue evolution at an isolated intersection cycle by cycle with constant arrival and continuous queue. Two objective functions are provided in their models. One is a quadratic form of the delay of each cycle during the whole oversaturated period; the other is a performance index, which is the combination of the total delay and stop penalty. This model is further extended to optimize an oversaturated signalized network (Chang and Sun 2004). In the extended model, the traffic propagation between adjacent intersections is modeled with the relations proposed by Isaksen and Payne (1973) and the offsets are calculated by the model proposed by Choi (1997).

The queue polygon approach can consider the link-blockage problem under the assumption that there are continuous queues, which means that those models would not work well for the transition period from undersaturated condition to oversaturated conditions. Furthermore, this method would have difficulty in analyzing lane-blockage.

Platoon dispersion model (Transyt)

The original platoon dispersion model is not suitable for oversaturated traffic conditions. Hadi and Wallace (1995) propose enhancements to Transyt-7F to optimize signal-timing plans under congested conditions. Transyt-7F release 8 (Li and Gan 1999) explicitly models the link-spillback and lane-blockage by reducing the corresponding link saturation flow. Park et al. (1999) employ the queue polygon method to compute the queue delay, and tracks the link-blockage by continuously checking the end-of-queue vehicle. The resulting model is solved with their GA, which is enhanced in a late study (Park et al. 2000). Along this research line, Ceylan et al. (2010) propose a total delay estimation model based on Transyt traffic model, which consists of a uniform component and a random oversaturation component, and solve it with a Quasi-Newton method.

Store-and-forward based approach

The store-and-forward model of traffic network is proposed by Gazis and Potts (1963) for oversaturated intersections. The basic equation of this store-and-forward model is:

$$u_i(k) = s_i g_i(k)/c \qquad\qquad (2\text{-}3)$$

in which $u_i(k)$ is the outflow of link i at time k; $g_i(k)$ is the green time for the target stream at time k; s_i is the corresponding saturation flow, and; c is the cycle length of the signal. Equation (2-3) describes the traffic flow process without use of discrete variables. The main idea is using the average flow of the whole cycle to represent the actual flow during the green time, i.e., it assumes there is a continuous outflow from each network link and there is sufficient demand.

Through this simplification, this store-and-forward model enables the control model to use highly efficient global optimization algorithms with polynomial complexity, such as linear

programming and quadratic programming, which allows coordinated control of large networks in real time. However, the store-and-forward model is unable to provide more accurate representation of traffic dynamics.

Cell-Transmission based approach

The Cell Transmission Model (CTM) (Daganzo 1994; Daganzo 1995) is a finite difference approximation of the traffic flow model (LWR model) by Lighthill and Whitham (1955) and Richards (1956). Its core concept is to divide the target roadway into homogeneous sections (cells), whose lengths are equal the distance traveled by a vehicle in free flow speed during one unit interval. CTM is capable of replicating kinematic waves, queue formation and dissipation. Lin and Ahanotu (1995) perform a validation for the CTM with respect to the formation and dissipation of queues in the content of first-order characteristics.

CTM is employed by researchers to study signal optimization (Lo 1999; Lo et al. 2001; Lo and Chow 2004) and promising results are reported. Ziliaskopoulos (2000) uses this model to study the system optimum dynamic traffic assignment problem with single destination and linear link cost functions. The most recent Transyt (Binning et al. 2008) by TRL integrates CTM as an alternative system performance index to the previous Platoon Dispersion Model (PDM). By using CTM, Transyt can consider the link-blocking and time-varying flow analysis.

Despite the promising results from those models, some critical issues remain un-addressed. First, most studies model the dynamic queue evolution either at a link-based level or at an individual movement-based level, which could result in difficulty in integrating with complex signal phases. Second, the lane-blockages, which are very common during congested conditions, have not been explicitly and dynamically modeled. Although some researchers have

30

attempted to address these issues by developing mesoscopic or microscopic traffic-simulation-based signal optimizers (Park et al. 1999; Stevanovic et al. 2007; Yun and Park 2006), there are still potential difficulties with these mesoscopic or microscopic simulation models. First, concerns are often raised regarding the computing efficiency and efforts needed to calibrate various behavioral parameters for microscopic-simulation methods. Secondly, the assumptions of microscopic simulation are difficult to validate because human behavior in real traffic is difficult to observe and measure.

2.8. *Conclusions*

This chapter summarizes the existing models related to the integrated control of freeways and arterials. Those models have been classed into different groups based on the features of their traffic flow models. The findings from the literature indicate awareness by researchers that pushing the problem from arterials to freeways or vice versa could not solve the congestion problem. The congestion problem around a freeway interchange could not be solved either on the freeways or on the arterials side. To eliminate this congestion, we should balance the delays of freeways and arterials and improve the overall system performance instead of individual subsystem performance.

As summarized in the literature review, several models have been proposed control arterials and freeways jointly with various purposes. Those models consider system-wide performance instead of individual subsystem performance, and provide integrated control for freeway and arterial systems. However, none of the existing models has considered the interaction between freeways and arterials, especially off-ramp queue spillback problem, which is a reasonable approximation for under-saturated traffic conditions but could not properly

represent the traffic dynamics under oversaturated traffic conditions. Integrated controls could not be achieved by just putting freeways and arterials in the same model. The major purpose of integrated control is to solve the interface problems between freeways and arterials. We have discussed in detail the potential issues around a congested interchange in Chapter 1. In those discussions, we have concluded that the interface problems which include on-ramp and off-ramp queue spillback could not be solved in either system. Without properly capturing the traffic dynamics between freeways and arterials, an integrated control model could produce even worse results than the separate control models. The on-ramp and, especially, the off-ramp spillback problems have not attracted sufficient research interests. Those research gaps provide an opportunity for this dissertation to make its own contribution.

This dissertation proposes an integrated control model for freeway interchanges to balance the delay of freeways and arterials and achieve a system optimal. To achieve this goal, this dissertation develops traffic flow models based on the Cell-Transmission concept to capture the lane-blockage, on-ramp spillback, and off-ramp spillback simultaneously. With those models, it is expected that the proposed model can optimize the arterial signal timings and on-ramp metering to obtain a near optimal system-wide solution.

Chapter-3: System Framework and Primary Tasks

3.1. *Introduction*

This chapter presents the overall structure of the proposed integrated control model for managing recurrent congestion of freeway interchange. The interrelations between its principal components, along with critical control factors and underlying assumptions, constitute the core of this chapter.

The remaining sections of this chapter are organized as follows: Section 3.2 presents the major research issues and challenges involved in developing such a system for dealing with recurrent congestion at freeway interchanges, including off-ramp overflow, on-ramp queue spillback, and intersection lane blockage. Section 3.3 presents the control flowchart of the proposed integrated control system, based on the research scope and intended applications. Section 3.4 describes the functions for all principal control component and their operational interrelations.

3.2. *Key Research Issues*

The proposed integrated control system for managing recurrent congestion aims to maximize the operational efficiency for target freeways and arterials. Based on the research objectives and the required system features stated in Chapter 1, some major research issues to be addressed in developing an integrated control system are listed below:

- How should the lane- and link-blockage traffic patterns and their evolution from moderate congestion to oversaturated conditions be modeled?

- How should the complex interrelationships between traffic queue at the on-ramp, off-ramp, and freeway mainline segment under various congestion levels be modeled?

- How should we balance the delay between freeway and arterial so as to achieve a system-wide optimal state?

- How should we formulate an integrated control model that can effectively and reliably account for complex interrelations between freeway ramp and arterial traffic flows, and can yield effective solutions for real-time applications?

To resolve the above research issues, the research work has been organized into the following tasks:

Task 1: Modeling the interrelations between the off-ramp queue and its neighboring intersection lane-blockages as well as on-ramp spillback under various congestion levels. This is needed to properly consider the recurrent congestion patterns where saturated local traffic conditions may cause the formation of off-ramp queues, but not affect the operational capacity of the mainline segments;

Task 2: Formulating an integrated control model for freeway interchanges that can trade-off between delays on arterials and freeways, and providing a system-wide near optimal control solution. The proposed model is expected to concurrently optimize the on-ramp and off-ramp controls as well as signal timings on the connected local arterials;

Task 3: Developing a generalized interchange control model for recurrent congestion scenarios in which an off-ramp queue may spillback to the freeway mainline and interfere with the upstream merging traffic flow from upstream ramps. The focus of this task is to tackle the severe congestion where both the freeway and local arterial are oversaturated, and the off-ramp queue may significantly reduce the freeway through capacity and spill back to its upstream ramp;

Task 4: Designing efficient solution algorithms for both the base model for off-ramp control and the extended model for integrated control of freeway interchagnes. The proposed algorithm shall have the capacity of generating efficient control parameters in response to the information deficiencies and dynamic traffic flow interactions under various congestion levels;

Task 5: Evaluating the effectiveness of the proposed models with numerical experiments. The primary focus of this task is to ensure the applicability of the proposed control models to the traffic system that often experiences off-ramp queue spillback during peak hours.

3.3. *System Control Structure*

To ensure that the products from each of the above tasks can be integrated into a seamless control structure and activated based on the detected congestion level, the overall control architecture for the proposed system is organized into the following three levels:

Level-1: **The off-ramp queue spillback to the freeway mainline**: The control model should consider the delay on the freeways in optimizing the signal timings at the off-ramp and intersections in the connected arterials.

Level-2: **The on-ramp queue spillback to its upstream intersection**: An insufficient metering rate may cause the on-ramp vehicles to block one or more arterial through lane(s), and consequently block the through traffic to spill back to its upstream intersections if the arterial through demand exceeds its remaining capacity. The control model should activate its oversaturated intersection module to maximize the total throughput within the control boundaries.

Level-3: **The freeway mainline at the interchange area experiences moving queue which spills back to its upstream interchange.** In this scenario, both freeways and arterials

have reached their capacity, and the on-ramp queue has spilled back to its neighboring arterial through lanes, while the off-ramp queue has propagated to its upstream interchange.

Figure 3-1 illustrates the feedback operating structure of the proposed integrated control system for freeway interchanges. The system takes traffic demand and the existing signal timing as input, and then executes the simulation model to check whether the queue spills back at on-ramps and off-ramps. The on-ramp metering and arterial signals (including the off-ramp signals) are operated independently if neither experiences any queue spillback. If the queue spillback only occurs at off-ramps, the system activates the off-ramp integrated control model to maximize the system performance. Likewise, if only the on-ramp queue spills back to arterials, the proposed system activates only the on-ramp integrated control model to balance the freeway and arterial delays. All models in the proposed integrated system are activated if both on-ramps and off-ramps suffer from queue spillbacks. The simulation module also checks the freeway mainline spillback, and executes the multi-interchange model to coordinate all control plans activated at those two neighboring interchanges.

Note that the entire system illustrated in Figure 3-1 requires various inputs for its on-line operations, which include:

- **The roadway geometric features,** such as the number of ramps, and distance between ramps and intersections, the length for left-turn bay, deceleration lane, and on-ramp acceleration lane.

- **Traffic volumes** on the freeway and arterial mainlines, ramps, and intersections;

- **Turning proportions** at both neighboring intersections and off-ramps.

- **Operational constraints for signal timing and metering plans;** and

- **Traffic flow parameters** that reflect local driving characteristics.

Figure 3-1 System control flowchart

3.4. *Principal System Modules and Key Functions*

To provide the aforementioned operational functions in response to various levels of saturated and oversaturated traffic congestion, the proposed integrated control system has the following modules: an arterial signal timing optimization module, an off-ramp integrated control module, an on-ramp integrated control module, a single-interchange integrated control module, and a multi-interchange integrated control module.
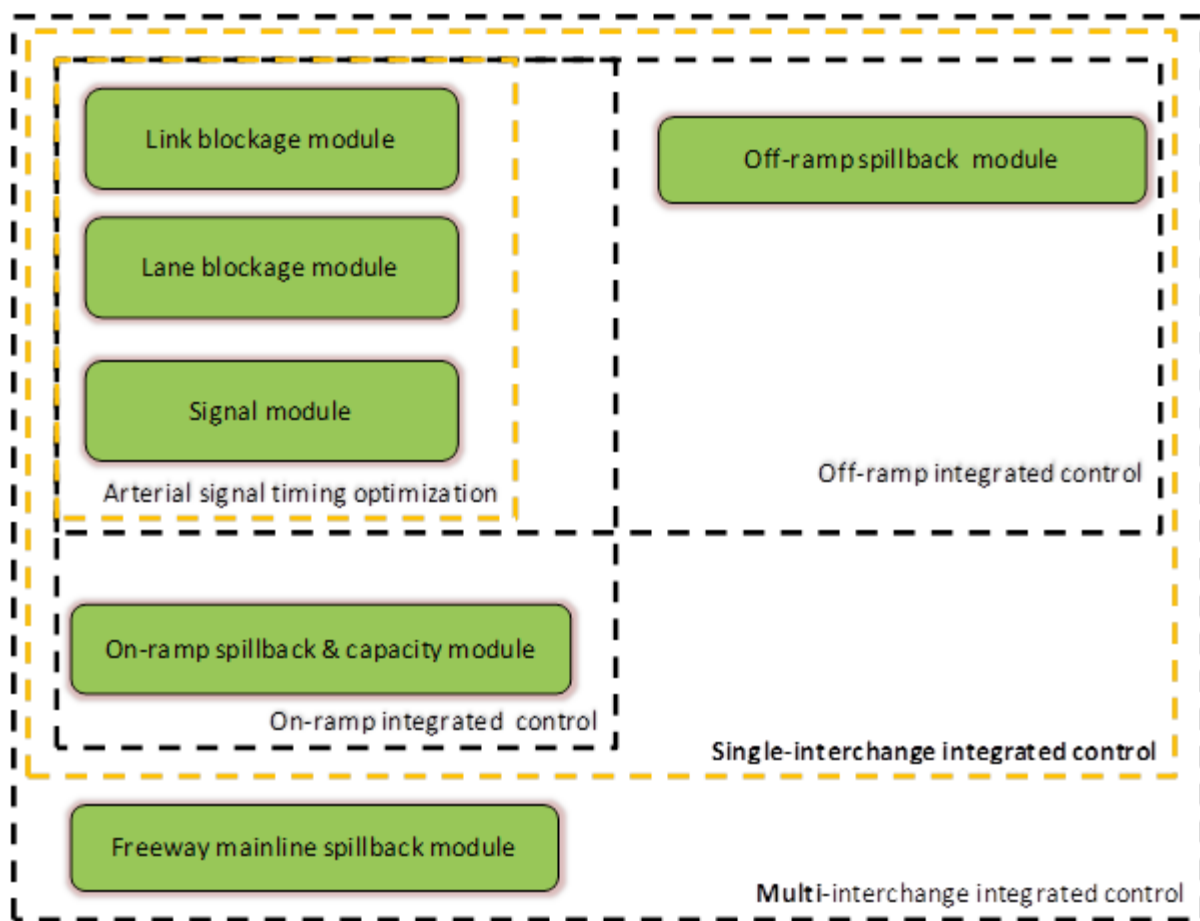
Figure 3-2 Key system modules

The interrelations between those modules are illustrated in Figure 3-2. Note that the

Arterial signal timing optimization module aims to optimize the cycle length, offset, and green split for all signals within the control boundaries under both under-saturated and oversaturated

conditions. The proposed module should be able to take into account the lane-blockage between lanes and spillback between intersections. The **Off-ramp integrated control module** is designed to incorporate the ramp queue delay and its impact on the freeway mainline flow in the arterial signal optimization module. The **On-ramp integrated control module** is designed to extend the function of the local arterial signal optimization module so that it can concurrently account for the on-ramp metering in the system-wide signal control problem. The **Single-interchange integrated control module** integrates all modules for on-ramp, off-ramp, and signal optimization into an interchange-wide operation system. The **Multi-interchange integrated control module** is designed to function as a coordination control model, which can concurrently account for both freeway and arterial traffic conditions at neighboring interchanges and produce system-wide control plans.

Chapter-4: Modeling Arterial Signal Optimization with Enhanced Cell Transmission Formulations

4.1. *Introduction*

This chapter addresses the critical lane-blockage issue in arterial signal optimization with enhanced formulations for the CTM diverging model, which corresponds to Task 1. The proposed model takes full account of the lane channelization effects on turning traffic, and captures the lane-blockage among different movements. Based on the enhanced formulations for lane-blockage, this chapter will present an arterial signal optimization model that can account for oversaturated conditions at oversaturated intersections.

The rest of this chapter is organized as follows: Section 4.2 presents the modeling methodology for arterial traffic dynamics under oversaturated traffic conditions. Section 4.3 presents the signal optimization model and a GA-based solution algorithm. A brief summary is included in Section 4.4.

4.2. *Modeling Methodology for Arterial*

To model the temporal and spatial interactions of traffic flows at an arterial intersection, we can conceptually divide each approach into the following four zones: merging, propagation, diverging, and departure zones (see Figure 4-1). Vehicles entering such a link will move over these four zones and then diverge to their respective destinations. During congested peak hours, left-turn and through vehicles in the same approach may block each other due to spillback if the bay length is insufficient or signal timings are not properly designed in response to the traffic

demand. The queue caused by lane-blockages may then spill back to the upstream intersections under saturated traffic conditions.
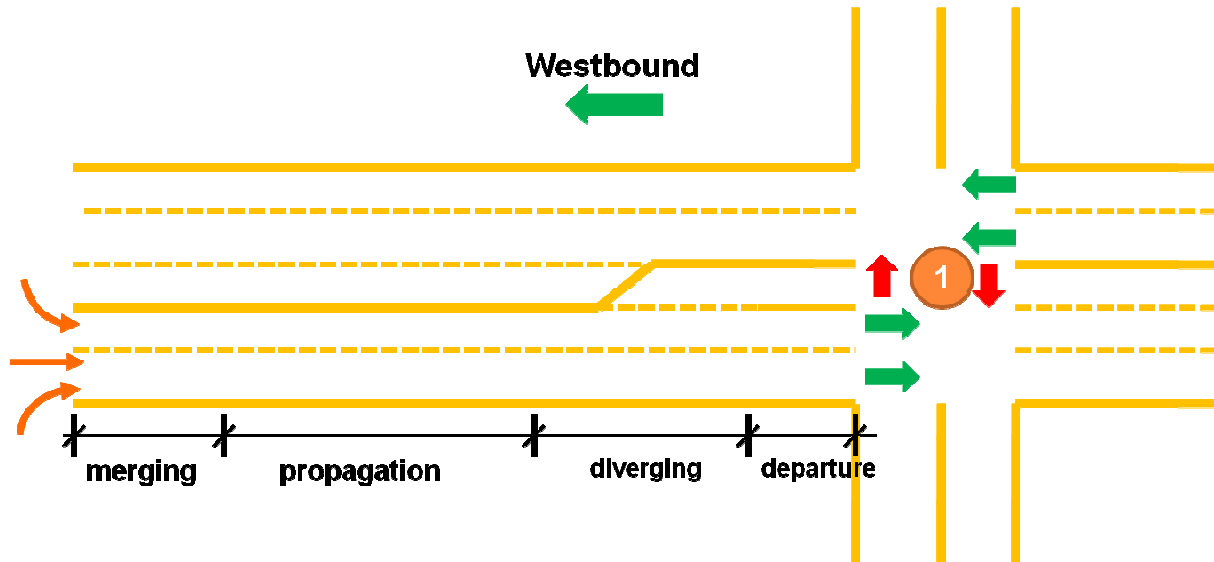
Figure 4-1 Traffic dynamic of a signalized intersection approach

To the optimize signal times for arterials experiencing lane-blockage at some intersection, this study first employs the Cell Transmission concept to formulate the flow interactions in the above four zones. The Cell Transmission Model (CTM) (Daganzo 1994; Daganzo 1995) is a finite difference approximation of the traffic flow model developed by Lighthill and Whitham (1955) and Richards (Richards 1956). Its core concept is to divide the target roadway into homogeneous sections (cells), whose lengths equal to the distance traveled by a vehicle in the free flow speed during one unit interval. CTM assumes that the relationship between traffic flow (q) and density (k) is of the form depicted in Figure 4-2.
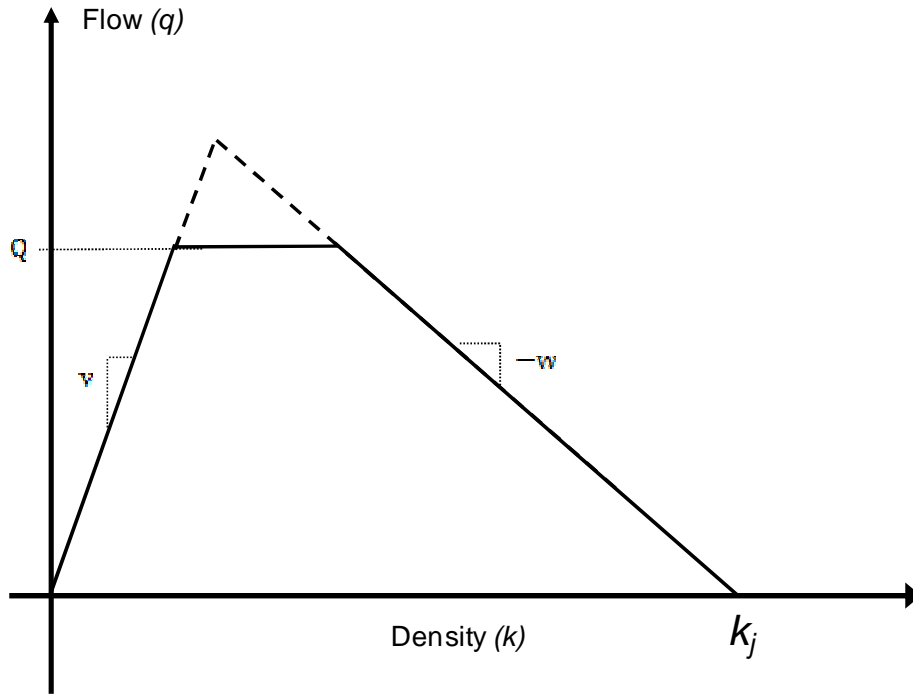
Figure 4-2 The density-flow relationship of trapezoid shape

The states of the traffic system at any time instant are tracked using the number of vehicles in each cell, denoted as n_i^t. In addition to the number of vehicles, the following parameters are commonly used in the CTM model, where time period t represents the time interval $[t, t + 1]$:

- N_i^t is the buffer capacity, defined as the maximum number of vehicles that can present in cell i at time t, which is the cell length multiplied by the jam density;

- Q_i^t is the flow capacity in time t, and defined as the maximum number of vehicles that can flow into cell i, which can be computed as the cell's saturated flow rate multiplied by the time interval duration;

- y_{ij}^t is defined as the number of vehicles leaving cell i and entering cell j during period t.

There are three types of cells defined in the CTM model: the ordinary cell, the merging cell and the diverging cell. The ordinary cell has just one upstream cell and one downstream cell;

43

the merging cell has more than one upstream cell and one downstream cell; the diverging cell has only one upstream cell and more than one downstream cell. The recursive relation of the CTM model can be expressed as follows:

$$n_i^{t+1} = n_i^t + y_{i,in}^t + y_{i,out}^t \tag{4-1}$$

$$y_{i,in}^t = \sum_{k \in \Gamma(i)} y_{ki}^t, \; and \; y_{i,out}^t = \sum_{j \in \Gamma^-(i)} y_{ij}^t \tag{4-2}$$

Equation (4-1) represents the flow conservation relationship at the cell level, which means that the number of vehicles in a cell in the next time interval equals the number of vehicles in this interval plus the difference between all entering and departing vehicles. The second and third terms in Equation (4-1) will vary with the cell category, where y_{ij}^t needs to be computed with a traffic flow-density relationship. The following sections will detail how the core CTM concept can be applied in formulating traffic flow interactions in these four identified traffic zones.

To represent complex traffic behavior such as lane-blockage, it is necessary to track the number of vehicles for each movement. Therefore this study employs the following equations to track the number of vehicles in each movement:

$$n_{i,m}^{t+1} = n_{i,m}^t + y_{i,m,in}^t - y_{i,m,out}^t \tag{4-3}$$

where $n_{i,m}^t$ is the number of vehicles for movement m of Cell i, and $y_{i,m,in}^t$ is the number of those vehicles that travel from upstream cell(s) to Cell i and stay in the movement m of Cell i, and $y_{i,m,out}^t$ is the number of vehicles that leave movement m of Cell i during period t .

4.2.1 Merging zone

In the merging zone, the vehicles from different upstream approaches will join and form a single traffic stream. During oversaturated traffic conditions, the queue can spillback and block the upstream traffic as shown in Figure 4-3.
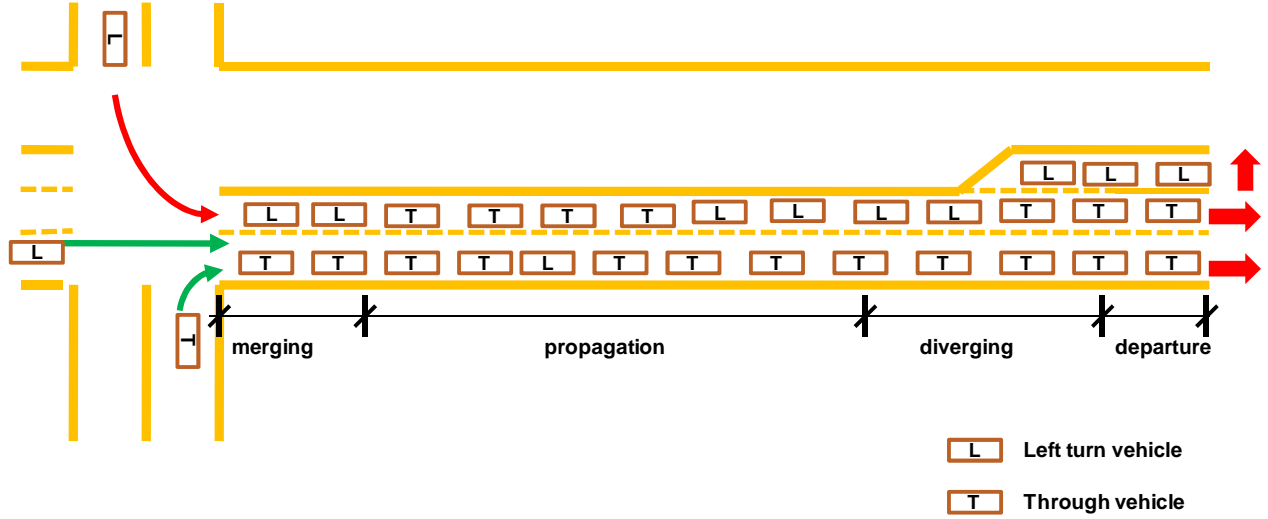
Figure 4-3 Link spillback blockage at merging zone

The merging cell is designed for modeling the traffic flow interactions in the merging zone. As illustrated in Figure 4-4, Cell C represents the merging zone; Cells A, B, D represent the upstream through, right-turn and left-turn approaches respectively. At signalized intersections, since the entering traffic streams are given different priorities to enter the merging zone based on the signal phasing plan, we can then use Equations (4-4) to capture such cell-level relations.

$$y_{iC}^t = \min\{n_i^t, Q_i^t, \delta[N^t{}_C - n_C^t]\}, i = A, B, D \qquad (4\text{-}4)$$

where $\delta = 1$, if $n_i^t \leq Q_i^t$, and $\delta = \frac{w}{v}$, if $n_i^t > Q_i^t$, in which w represents the backward propagating speed of the disturbances and v is the free flow speed. When the merging zone represented by cell C is full (i.e., the number of vehicles in cell C, n_C^t, equal to its buffer capacity, N_C^t), no vehicle can enter the merging zone (i.e., $N_C^t - n_C^t = 0$ which implies $y_{iC}^t = 0$).
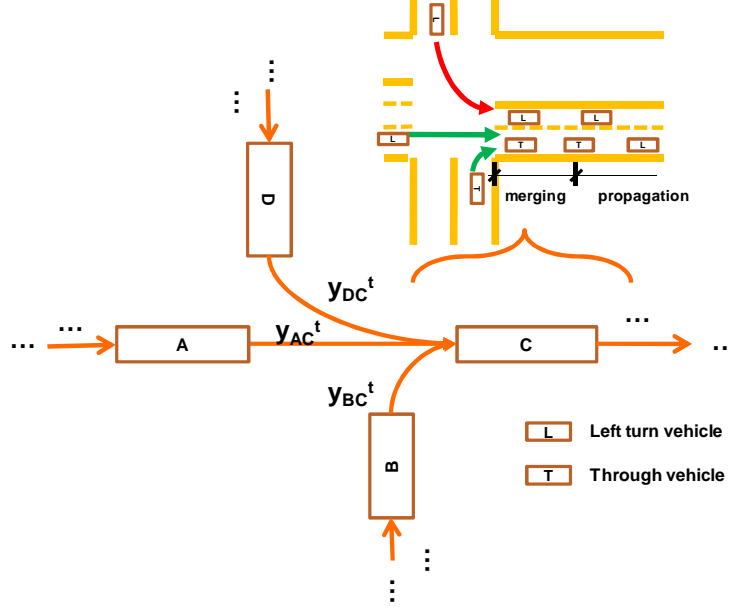
Figure 4-4 The merging zone represented by a merging cell

To represent the flow relation for each movement, this study employs the turning proportion method which means that when a platoon of vehicles arrives at a link, the proportion of vehicles in each movement will be assumed to be known and denoted by r_{lm}^t, where l, and m represent the link identity number and movement, respectively. Following this assumption, the number of vehicles for each movement can be updated as in equation (4-5).

$$y_{C,m,in}^t = r_{lm}^t(\sum_i y_{ic}^t), i = A, B, D \ for \ each \ movement \ m \tag{4-5}$$

4.2.2 Propagation zone

In the propagation zone, the interactions between vehicles increase with traffic volume, which then reduces the traffic speed. From the aggregate perspective, the flow-density relation can represent this effect. Hence, to compute the optimal signal plan for an arterial, we need to

formulate the temporal and spatial relations of traffic over the links between adjacent

intersections.
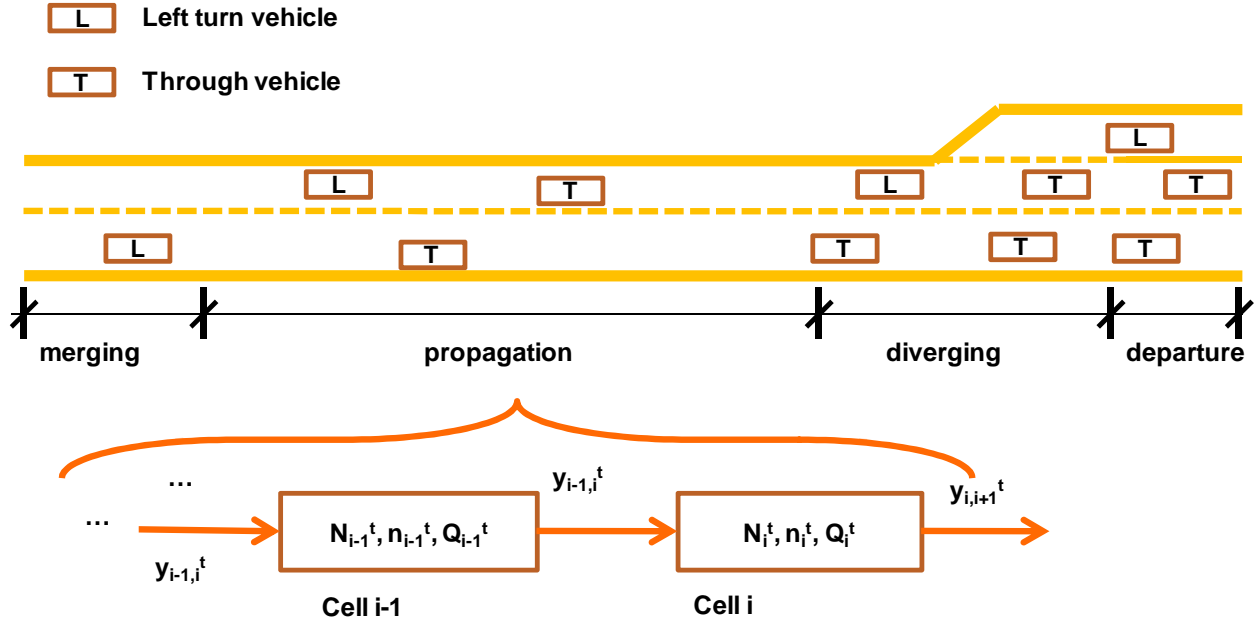
Figure 4-5 The propagation zone represented with ordinary cells

For such needs, this study employs the ordinary cell to capture these vehicle interactions

in the propagation zone. As illustrated in Figure 4-5, the number of cells in the propagation zone

may vary with the link length. Each ordinary cell has one upstream cell and one downstream cell.

The number of vehicles which can exit cell i and enter cell $i+1$ in time t ($y_{i,i+1}^t$) can be

determined with Equation (4-6), a simplified flow-density relation proposed by Daganzo (1956)

that can capture the traffic dynamics under various traffic conditions:

$$y_{i,out}^t = y_{i+1,in}^t = \min\{n_i^t, Q_i^t, \delta[N_{i+1}^t - n_{i+1}^t]\} \tag{4-6}$$

If we define $S_i^t (= min\{Q_i^t, n_i^t\})$ as the sending capacity, and $R_i^t (= \min\{Q_i^t, \delta(N_i^t - n_i^t)\})$

as the receiving capacity of cell i, then Equation (4-6) can be restated as Equation (4-7):

$$y_{i,i+1}^t = \min\{S_i^t, R_{i+1}^t\} \tag{4-7}$$

$$y_{i,m,out}^t = y_{i,out}^t \times \frac{n_{i,m}^t}{n_i^t} \; for \; each \; movement \; m \tag{4-8}$$

4.2.3 Diverging zone

In the diverging zone, vehicles bound to different destinations may join different queues. Under oversaturated conditions, blockage between different movements could occur. For instance, depending on the bay length, the left-turn queue could spill back and block the through traffic. For convenience of illustrating the modeling concept, let us consider only the interactions between left-turn and through vehicles. However, the concepts presented in this section can be extended to other types of lane blockage. An intersection approach with left-turn and through lanes could have two possible types of lane blockage as shown in Figure 4-6 and Figure 4-7, respectively.
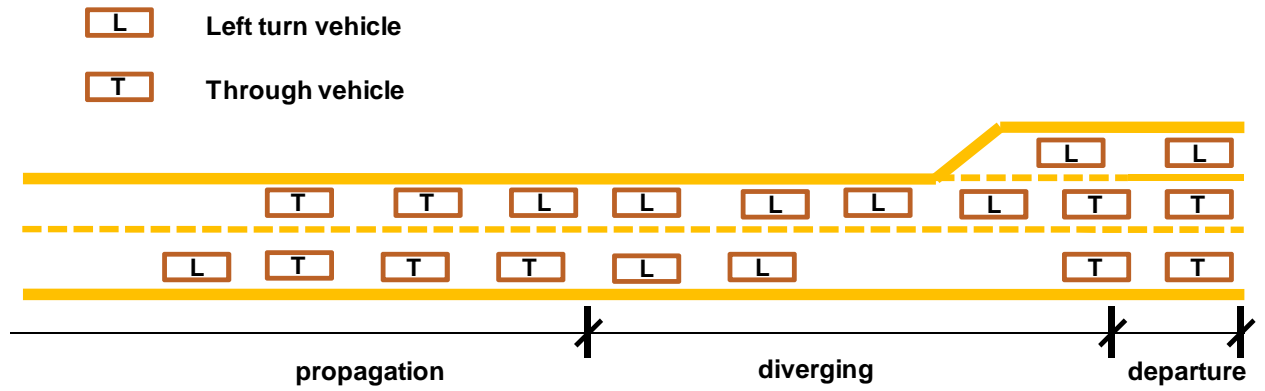
Figure 4-6 Left-turn blocks through traffic
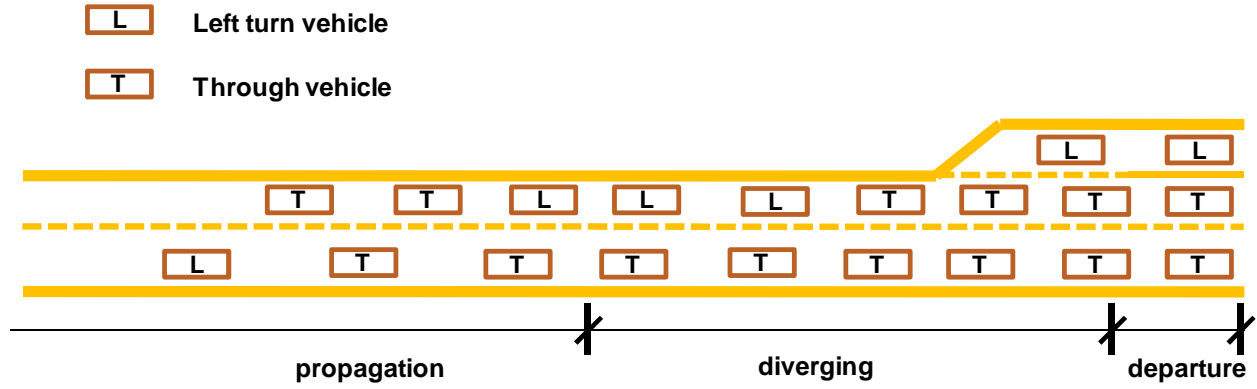
Figure 4-7 Through blocks left-turn traffic

The diverging movements in Figure 4-6 and Figure 4-7 are typically modeled with a diverging cell in the literature (1994) since there exist multiple exit movements. However, the traditional CTM diverging cell does not consider the blockage effect between lanes, which is quite common under over-saturated conditions. To realistically capture the queue and blockage effect between neighboring movements, this study proposes an enhanced diverging model that employs the sub-cell concept to represent each type of movement.
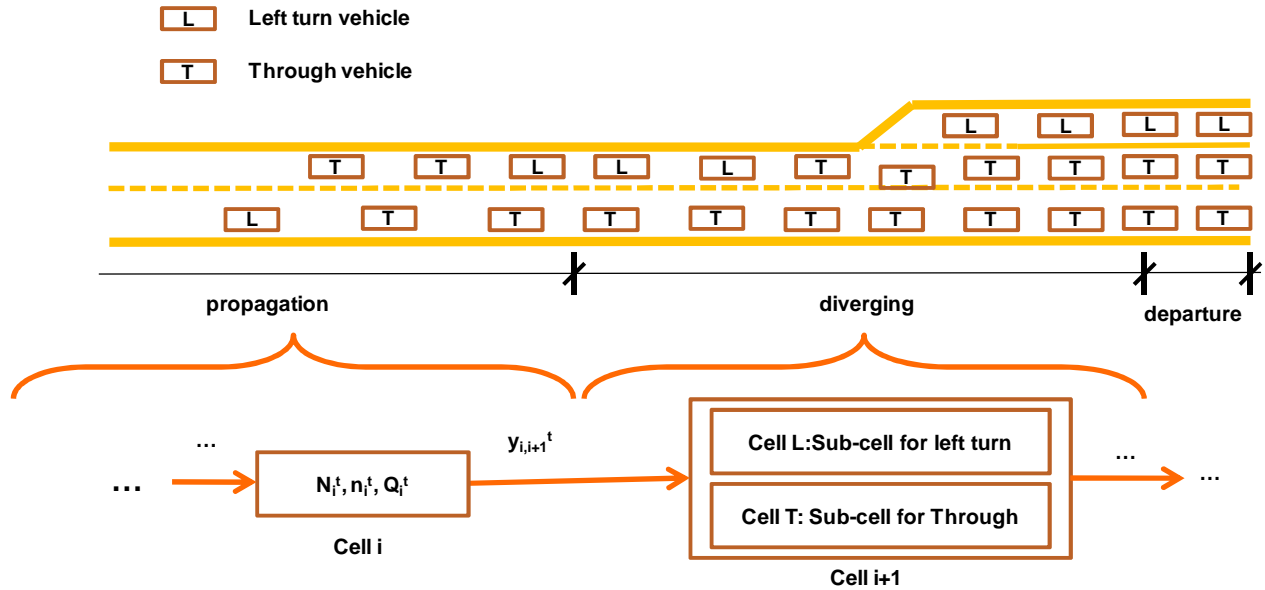
Figure 4-8 The illustration of sub-cell concept

As shown in Figure 4-8, the diverging zone link is presented with a diverging cell, Cell *i+1*, which is further divided into two sub-cells, sub-cell L for left-turning and sub-cell T for through traffic.
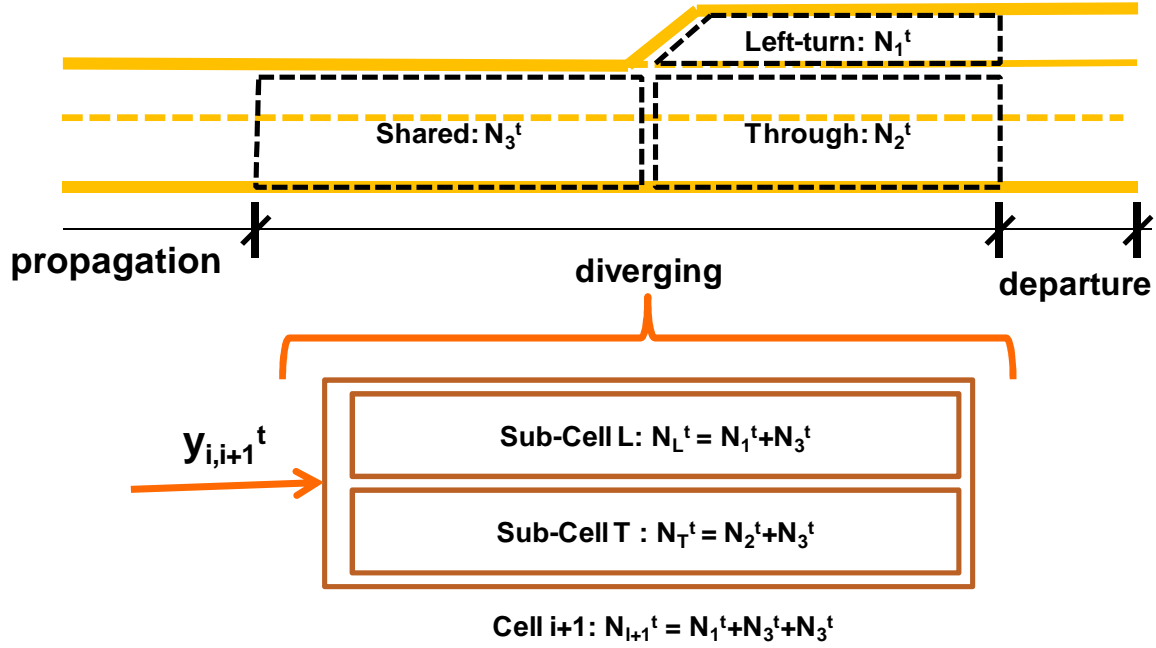
Figure 4-9 The sub-cell representation of a signalized diverging cell

The diverging zone can be further divided into the following three zones, as illustrated in Figure 4-9, in which Zone 1, denoted by N_1^t, is the space exclusively reserved for left-turn traffic; Zone 2, N_2^t, is the space used only for through traffic; and Zone 3, N_3^t, is the space shared by left-turn and through traffic. The buffer capacity of each sub-cell can be computed with Equations (4-9) and (4-10):

$$N_L^t = N_1^t + N_3^t \tag{4-9}$$

$$N_T^t = N_2^t + N_3^t \tag{4-10}$$

$$N_{i+1}^t = N_1^t + N_2^t + N_3^t \tag{4-11}$$

where Equation (4-11) captures the physical buffer capacity of the diverging cell $i+1$. We can divide these zones based on the channelization at a signalized approach. The buffer capacity of these sub cells explicitly reflects the turning bay effects. The flow capacity of each sub-cell can be computed with its lane number and the lane saturation flow rate.

Based on the above definitions, the status of these sub-cells can be modeled by the linear programming problem represented in Equations (4-12) to (4-17).

$$max \sum_m y_{i,m,out}^t \qquad (4\text{-}12)$$

$$\sum y_{i,m,out}^t \le R_{i+1}^t \qquad (4\text{-}13)$$

$$y_{i,m,out}^t \le \delta(N_{i,m}^t - n_{i+1,m}^t) \qquad (4\text{-}14)$$

$$\sum y_{i,m,out}^t \le S_i^t \qquad (4\text{-}15)$$

$$y_{i,m,out}^t \le S_i^t \times \frac{n_{i,m}^t}{n_i^t} \qquad (4\text{-}16)$$

$$y_{i,m,out}^t \le Q_{i,m}^t \qquad (4\text{-}17)$$

Equation (4-12) assumes that the traffic will try to fully utilize the available capacity and space. For instance, as depicted by Figure 4-10, when the left-turn queue spillback occurs, the left-turn vehicles may eventually occupy all the shared zone space if left-turn traffic continues to increase.
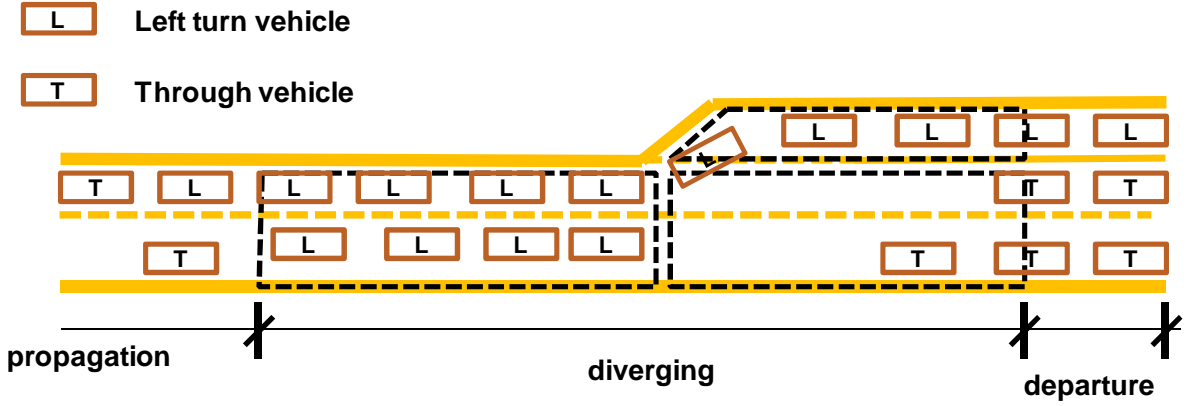
Figure 4-10 The Illustration of left-turn blocking through traffic

The new diverging model presented in this section offers the capability to explicitly model the effect of the turning bay, and capture lane-blockage as illustrated by Equations (4-12), (4-13), and (4-17). In the illustrative scenario, the third term in the parenthesis of Equation (4-12) will be the minimum of these three terms, which implies $w_i^t = R_L^t/r_L^t$ according to Equation (4-12). By substituting it into Equations (4-13) and (4-17), we can deduce that $y_{iL}^t = R_L^t$ and $y_{iT}^t = R_L^t \times r_T^t/r_L^t$. If R_L^t decreases, y_{iL}^t and y_{iT}^t will also decrease. When $R_L^t = 0$, it indicates that left-turn vehicle have blocked through traffic completely. We can perform the same analysis for the scenario of through blocking left-turn traffic.

4.2.4 Departure zone

The segment in the departure zone is modeled with a signalized cell. Its flow capacity at time t, Q_i^t is determined by the green time and defined as follows:

$$Q_i^t = Q_{i,max}\, g_i^t \tag{4-18}$$

where g_i^t is the green time in time interval t, which will be determined by Equations (4-33) and (4-34). Equation (4-18) enables the proposed model search those cycle lengths which are not multipliers of the period duration.

4.3. *An Optimization Model for Oversaturated Arterial Signals*

4.3.1 Objective functions

Depending on the traffic conditions, we can set the control objective function as maximizing the total system throughput or minimizing the total delay. Using the above formulations, the objective function of the proposed signal optimization model to maximize the system throughput can be expressed as:

$$Max \ (Throughput = \sum_{t=0}^{T}\sum_{j \in S} \sum_{i \in \Gamma^-(j)} y_{ij}^t) \qquad (4\text{-}19)$$

where S is the sink cell set, $\Gamma^-(j)$ is the upstream cell set of cell j, and T is total operation period.

In CTM, the length of each cell equals to the free-flow travel distance over a pre-specified unit of time, which means that the vehicles in each cell can either stay or move to the downstream cells. If we define the delay as the difference between a vehicle's actual travel time and its free speed travel time over a given distance(Hall 1993), the delay experienced by a vehicle in a cell can be computed by the time intervals in which it stays in the same cell. For instance, if a vehicle stays in the same cell over n consecutive unit intervals, then this implies that the vehicle has experienced n units of delay. More specifically, we can define the delay over each cell for time interval t as $d_i^t = \left(n_i^{t-1} - \sum_{j \in \Gamma(i)} y_{ij}^t\right) \times \tau$, where $\Gamma(i)$ the downstream cell is set of cell i and τ is the period duration. Thus, we can propose an alternative objective function of minimizing the total system delay as:

$$Min \ [total \ delay = \ \tau \sum_{t=0}^{T}\sum_{i}(n_i^t - \sum_{j \in \Gamma(i)} y_{ij}^t] \qquad (4\text{-}20)$$

As τ is a constant, the objective function of minimizing the system delay can further be

simplified as:

$$Min \: [Z = \sum_{t=0}^{T} \sum_{i} (n_i^t - \sum_{j \in \Gamma(i)} y_{ij}^t]$$ (4-21)

4.3.2 Signal timing operation
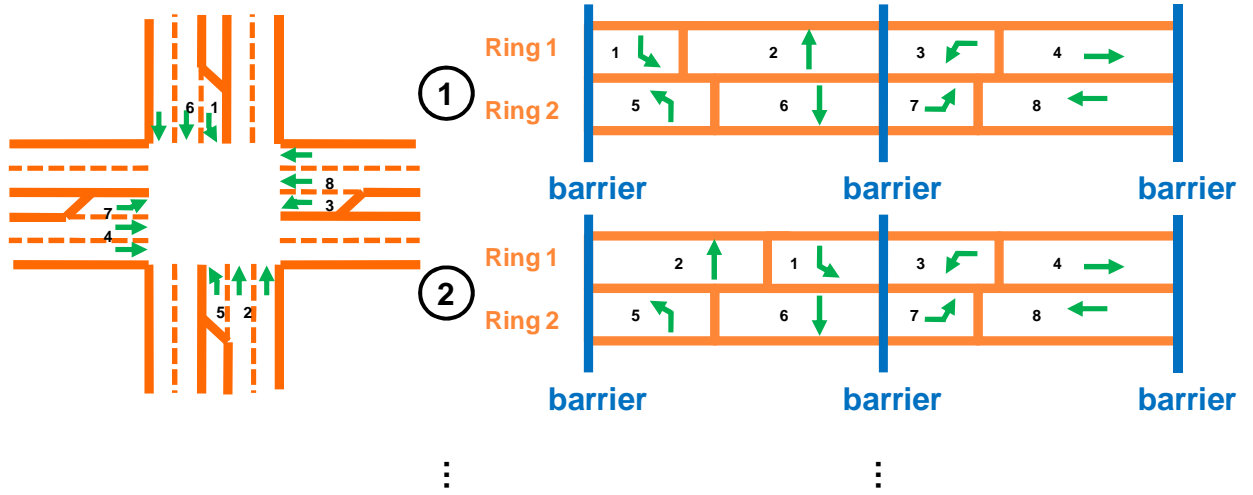
Figure 4-11 NEMA eight-phase signal timing structure

Figure 4-11 illustrates a typical four-leg intersection and the NEMA eight-phase structure.

The right-turn on red is assumed to be permitted in this study. The two-ring eight-phase structure

illustrated in Figure 4-11 can be represented with the following equations:

$$g_{k1} + g_{k2} = g_{k5} + g_{k6}$$ (4-22)

$$g_{k3} + g_{k4} = g_{k7} + g_{k8}$$ (4-23)

$$g_{k1} + g_{k2} + g_{k3} + g_{k4} = C_k$$ (4-24)

$$C_k = C/2^{h_k}$$ (4-25)

$$h_k = \begin{cases} 1, signal\ k\ has\ half\ common\ cycle\ length \\ 0, \qquad\qquad otherwise \end{cases} \qquad (4\text{-}26)$$

$$g_{kj} \geq MG_{kj}, j = 1, \dots, 8 \qquad (4\text{-}27)$$

$$MinC \leq C_k \leq MaxC \qquad (4\text{-}28)$$

$$0 \leq offset_k < C_k \qquad (4\text{-}29)$$

$$g_{kj}, C_k, offset_k\ are\ integers \qquad (4\text{-}30)$$

where g_{kj} is the green time for Phase j of signal k, C_k is the cycle length of signal k; MG_{kj} is the minimum green time of signal k phase j; $MinC$ is the minimum cycle length; $MaxC$ is the maximum cycle length; C is the common signal cycle length; h_k is a binary variable that indicates whether signal k has a half common cycle length or not; and t_k represents the offset of signal k. Equations (4-22) and (4-23) indicate the existence of the signal barrier. Equations (4-24) and (4-25) enforce the cycle length constraints. Equation (4-27) requires that the green time of each phase cannot be less than its minimal green time, and Equation (4-28) specifies a user-defined minimal and maximal cycle lengths. Equation (4-29) requires that the offset of signal k lie between 0 and its cycle length.

To compute the green time for each interval t of the departure cell, the green time of each phase should first be converted to time in a signal cycle:

$$G_{k0} = G_{k4} = 0; \ G_{ki} = \sum_{j=0}^{i-1} g_{kj}, \ for\ i = 1, 2, 3 \qquad (4\text{-}31)$$

$$G_{ki} = \sum_{j=4}^{i-4} g_{kj}, \ for\ i = 5, 6, 7 \qquad (4\text{-}32)$$

where G_{ki} is the green start time of phase i of signal k in its the signal cycle, as illustrated in Figure 4-11. If departure cell i is associated with signal phase j of signal k, the following equations will compute the green time of time interval t for cell i:

$$v_i^t = (t\tau + offset_k)\ mod\ C_k \tag{4-33}$$

$$g_i^t = \begin{cases} \max\{\min\{G_{k,j-1} + g_{kj}, v_i^t + \tau\} - \max\{G_{k,j-1}, v_i^t\}, 0\}, \\ \hspace{4cm} v_i^t + \tau \le C_k \\ \max\{min\{G_{k,j-1} + g_{kj}, C_k\} - max\{G_{k,j-1}, v_i^t\}, 0\} \\ + \max\{\min\{G_{k,j-1} + g_{kj}, v_i^t + \tau - C_k\} - \max\{G_{k,j-1}, 0\}, 0\}, \hspace{0.5cm} v_i^t + \tau > C_k \end{cases} \tag{4-34}$$

where v_i^t is the start time of time interval t in a signal cycle.

4.3.3 Solution Algorithm

In the proposed model, the decision variables are the cycle length, green time split, and the offset of each signal. This study proposes a Genetic-Algorithm-(GA)-based solution method for the proposed model. GA is a search technique based on the processes of natural selection and evolution. Recently, GA has been successfully applied to optimize signal timings under various traffic conditions (Ceylan 2006; Ceylan and Bell 2004b; Lo and Chow 2004; Park et al. 1999; Zhou et al. 2007).

The most critical part of developing a GA-based algorithm is to derive a good encoding scheme, i.e., how to represent possible solutions of the target problem by a gene series of 0-1 bits. This study employs an encoding scheme which includes the constraints (4-22)-(4-30), i.e., the signal timing decoded from the scheme will be feasible to constraints (4-22)-(4-30). The fraction-based decoding scheme, based on the NEMA phase's structure proposed by Park et al.(1999), can satisfy all the constraints except Equation (4-25). This study has enhanced this schema by including the half common cycle length for some signals with less traffic demands.

$$C = MinC + (MaxC-MinC) \times f_1$$

0   MinC   C   MaxC

$$C_k = C / 2^{I_k}$$

$$P_{k1} = max(MG_{k1}+MG_{k2},\ MG_{k5}+MG_{k6}) + [C_k - max(MG_{k1}+MG_{k2},\ MG_{k5}+MG_{k6}) - max(MG_{k3}+MG_{k4},\ MG_{k7}+MG_{k8})] \times f_2$$

P_{k1} $P_{k2}=C_k-P_{k1}$

$g_{k1}=MG_{k1}+(P_{k1}-MG_{k1}-MG_{k2})xf_3$ $g_{k3}=MG_{k2}+(P_{k2}-MG_{k3}-MG_{k4})xf_4$

g_{k1} $g_{k2}=P_{k1}-g_{k1}$ g_{k3} $g_{k4}=P_{k2}-g_{k3}$

MG_{k1} MG_{k2} MG_{k3} MG_{k4}

MG_{k5} MG_{k6} MG_{k7} MG_{k8}

g_{k5} $g_{k6}=P_{k1}-g_{k5}$ g_{k7} $g_{k8}=P_{k2}-g_{k7}$

$g_{k5}=MG_{k5}+(P_{k1}-MG_{k5}-MG_{k6})xf_5$ $g_{k7}=MG_{k7}+(P_{k2}-MG_{k7}-MG_{k8})xf_6$

$$Offset_k = C_k \times f_8$$
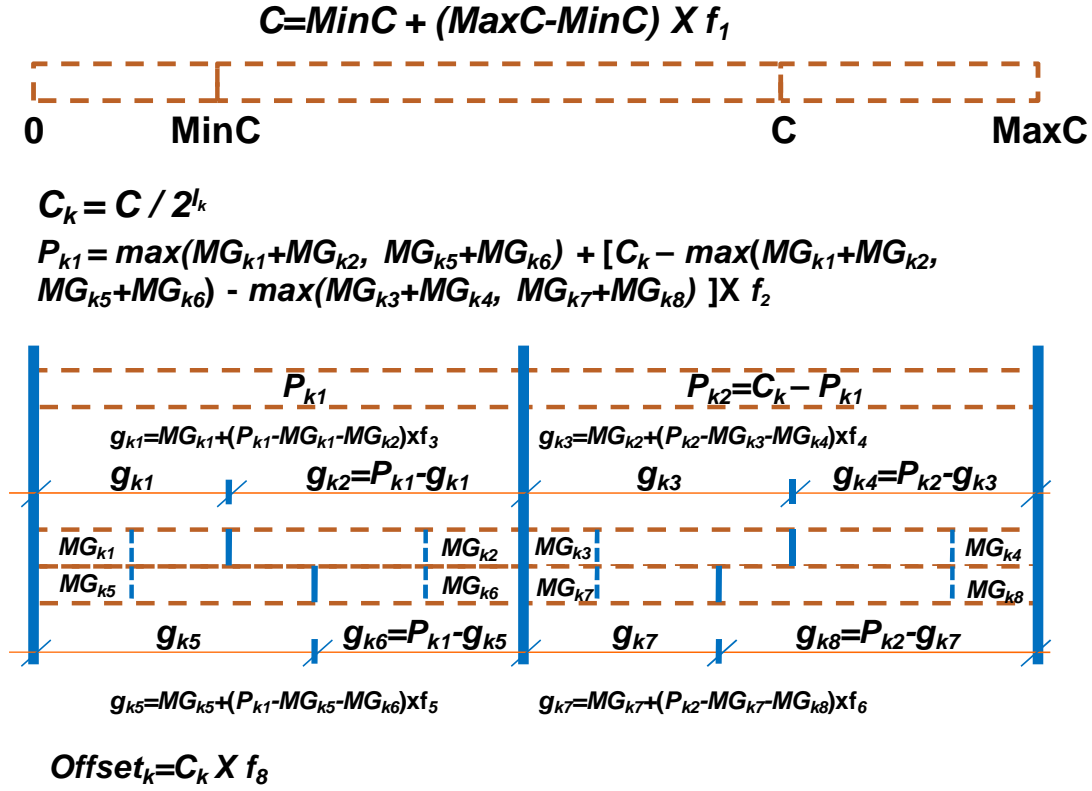
Figure 4-12 An enhanced fraction-based decoding scheme for signal timing

A detailed description of the original scheme can be found in Park et al. (1999). As illustrated in Figure 4-12, the proposed scheme sets the cycle length of signal *k* to half of a common cycle length if the half-cycle binary variable, I_k, is 1. Otherwise, the cycle length is set to be the full common cycle length.

4.4. *Summary*

This chapter has presented an enhanced Cell-Transmission Model for optimizing signal timings on congested arterials. The proposed model with its innovative sub-cell model is capable of capturing lane-blockage between neighboring lane groups due to queue spillback under high volume conditions. The signal optimization model presented here is designed to optimize the cycle length, split, and offset, under the presence of link-blockage and lane-blockage.

57

Chapter-5: An Integrated Single-interchange Control Model

5.1. *Introduction*

This chapter presents an integrated optimal control model. The proposed model is based on the arterial signal optimization model, but extends its control boundary to capture the impact around a congested interchange of off-ramp queue spillback to the through traffic of its upstream freeway. The inclusion of freeway mainline traffic delays caused directly and indirectly by the moving queue at the off-ramp allows the interchange control model to balance of congestion between freeways and arterials under oversaturated conditions.

Figure 5-1 illustrates a signalized interchange, which includes two closely spaced signals and two on- and two off-ramps. The distance between those two signals typically ranges from 500 ft in urban areas to 800 ft in suburbs. If a metering control strategy is implemented for the on-ramps, more traffic control devices are placed in such a tight area. The close placement of these control devices could incur off-ramp spillback, link-blockage, and lane-blockage, if those devices are not properly operated. Firstly, the short distance between the two signals greatly limits the queue storage, which increases the probability of queue spillback between them. The queue spillback could block the upstream approach if those two signals are not properly coordinated. Secondly, the on-ramp queue could spill back and block its upstream intersection. Thirdly, the off-ramp queue could spill back to its upstream freeway. When the exit volume from freeway exceeds the capacity of the connecting arterial, the exit queue eventually spills backs to its upstream freeway, and thus diminishes the through capacity of the freeway. This off-ramp spillback problem has been reported by researchers (Cassidy et al. 2002; Jia et al. 2004; Lovell 1997). The on-ramp and off-ramp queue spillback involve both freeways and arterials. To

address these issues, it is essential to balance the delays between freeways and arterials and

improve overall system performance instead of just emphasizing one of them.
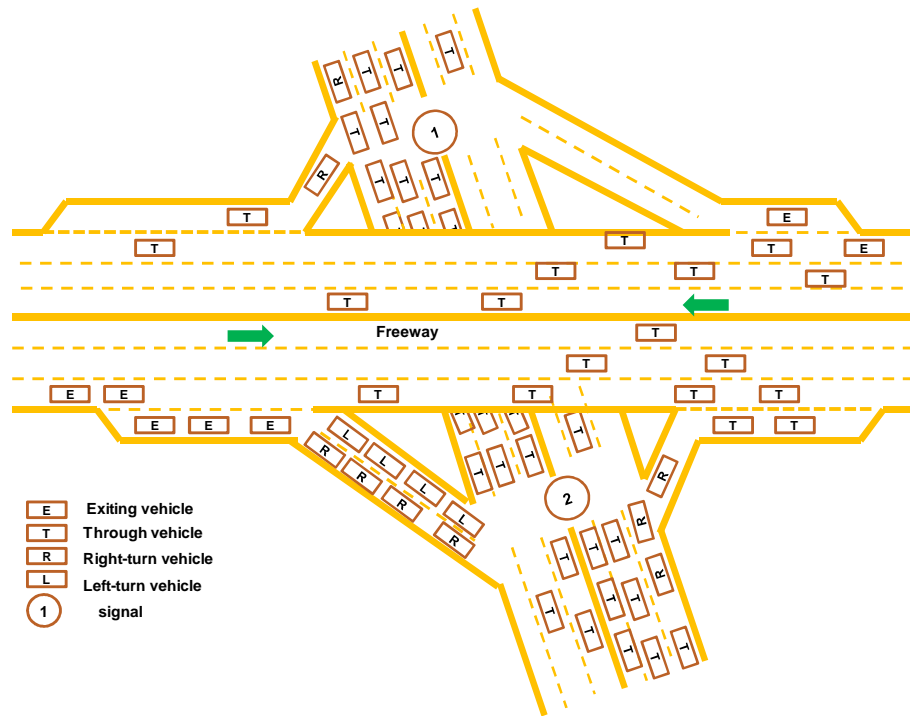
Figure 5-1 Graphical illustration of a signalized interchange

The focus of this chapter hereafter is to illustrate a freeway model component for

integration with the optimal arterial signal model. The proposed freeway traffic model can

capture the following two types of complex traffic flow interaction: (1) the impacts of arterial

traffic volume on the off-ramp queue length; and (2) the spillback of off-ramp queue on the delay

and operational capacity of its upstream freeway mainlines.

The remaining sections of this chapter are organized as follows. Section 5.2 presents the

core logic of the proposed freeway traffic flow model. Section 5.3 illustrates the mathematical

formulations for all complex traffic flow interactions and the solution algorithm. Section 5.4

summarizes the concluding comments.

5.2. *Modeling Methodology for Freeway*

As illustrated by Figure 5-2, a basic freeway mainline segment can be further divided into two different sections, A and B. Segment A (called the two-stream section) has two categories of vehicles. One travels to the downstream freeway mainline and the other to the downstream off-ramp. All vehicles of segment B (called the one-stream section) head to the downstream freeway mainline. The following subsections discuss how freeway traffic dynamics are modeled with the Cell-Transmission concept. The definitions of parameters are identical to those in Chapter-4.
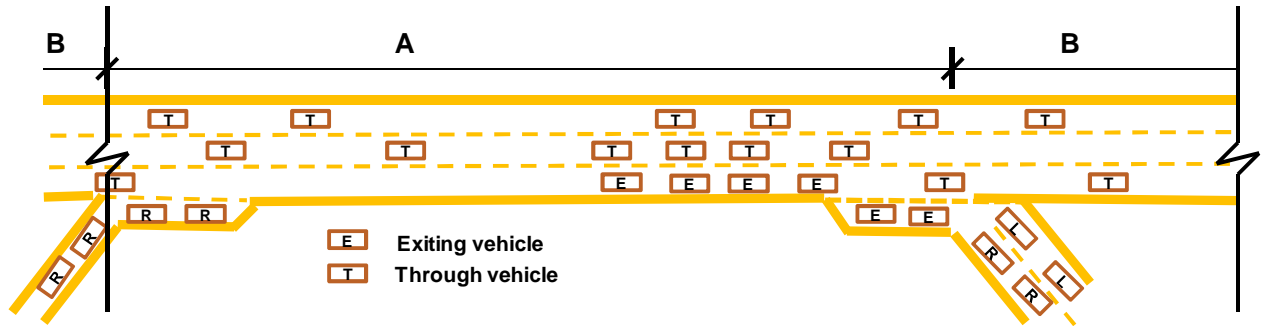
Figure 5-2 A basic freeway segment

5.2.1 Modeling of One-Stream Segments

The one-stream freeway section is the simplest case since each cell just has one entry cell and one exit cell as illustrated in Figure 5-3. The number of vehicles that can exit from Cell i and enter Cell $i + 1$ during time t, $y_{i,i+1}^t$, can be determined with Equation (5-1):

$$y_{i,i+1}^t = \min\{n_i^t, Q_i^t, \delta[N_{i+1}^t - n_{i+1}^t]\} \tag{5-1}$$

where $\delta = 1, if\ n_i^t \leq Q_i^t$, and $\delta = \frac{w}{v}, if\ n_i^t > Q_i^t$, in which w represents the speed with which disturbances propagate backward when traffic is congested, also known as the backward wave speed (Daganzo 1995), and v is the free-flow speed. If we define the sending capacity, $S_i^t (=$

61

$\min\{Q_i^t, n_i^t\}$), and receiving capacity $R_i^t (= \min\{Q_i^t, \delta(N_i^t - n_i^t)\})$, of Cell i, Equation (5-1) can be rewritten as:

$$y_{i,i+1}^t = \min\{S_i^t, R_{i+1}^t\} \qquad (5\text{-}2)$$
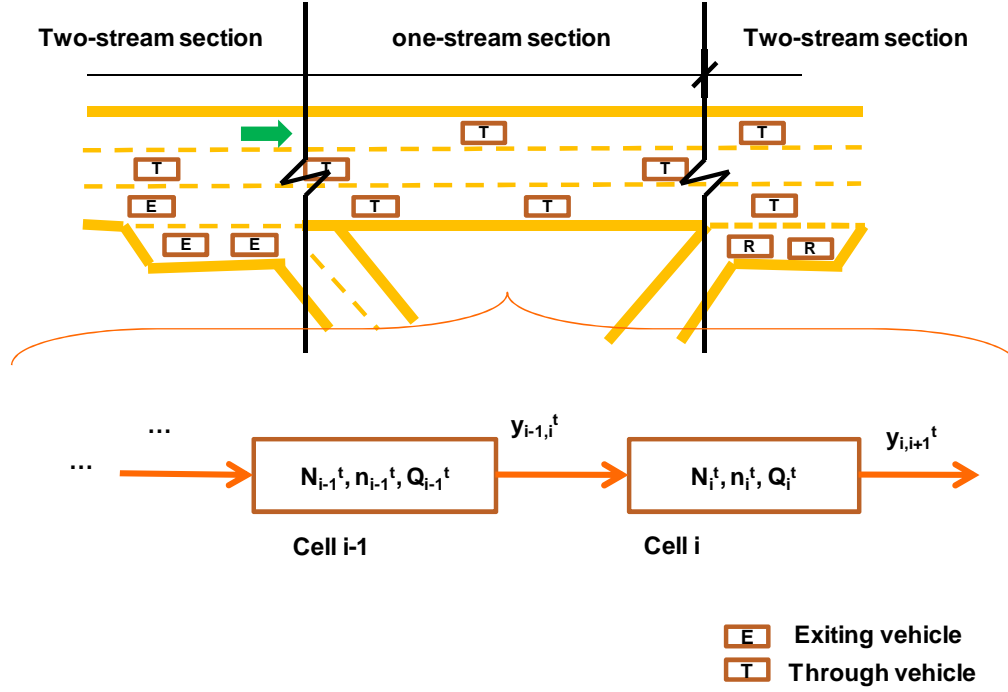
Figure 5-3 CTM modeling of traffic flow interactions in one-stream segment

Equation (5-3) represents the flow conservation relation at the cell level, which means that the number of vehicles in cell i at the beginning of the next interval, n_i^{t+1}, equals the number of vehicles of the current interval, n_i^t, plus the number of vehicles entering the cell, $y_{i-1,i}^t$, and minus the number of vehicles leaving it, $y_{i,i+1}^t$, during the current interval.

$$n_i^{t+1} = n_i^t + y_{i-1,i}^t - y_{i,i+1}^t \qquad (5\text{-}3)$$

5.2.2 Modeling of Two-Stream Segments

To model a Two-Stream freeway segment, it is necessary to add one more state variable, $n_{E,i}^t$, to track the number of exit vehicles, which can be computed with the following expression:

$$y_{E,i,i+1}^t = \min \{\gamma_{E,i}^t \times \min\{n_i^t, Q_i^t, \delta(N_{i+1}^t - n_{i+1}^t)\}, N_{E,i+1}^t - n_{E,i+1}^t\} \qquad (5\text{-}4)$$

where $y_{E,i,i+1}^t$ denotes the number of vehicles exit from Cell i to Cell $i+1$; $\gamma_{E,i}^t$ is the fraction of exit vehicles in Cell i for time interval t, which can be computed as $\gamma_{E,i}^t = n_{E,i}^t / n_i^t$; $N_{E,i+1}^t$ is the buffer capacity for exit traffic, i.e., the maximum number of exit vehicles that can be present in cell i.
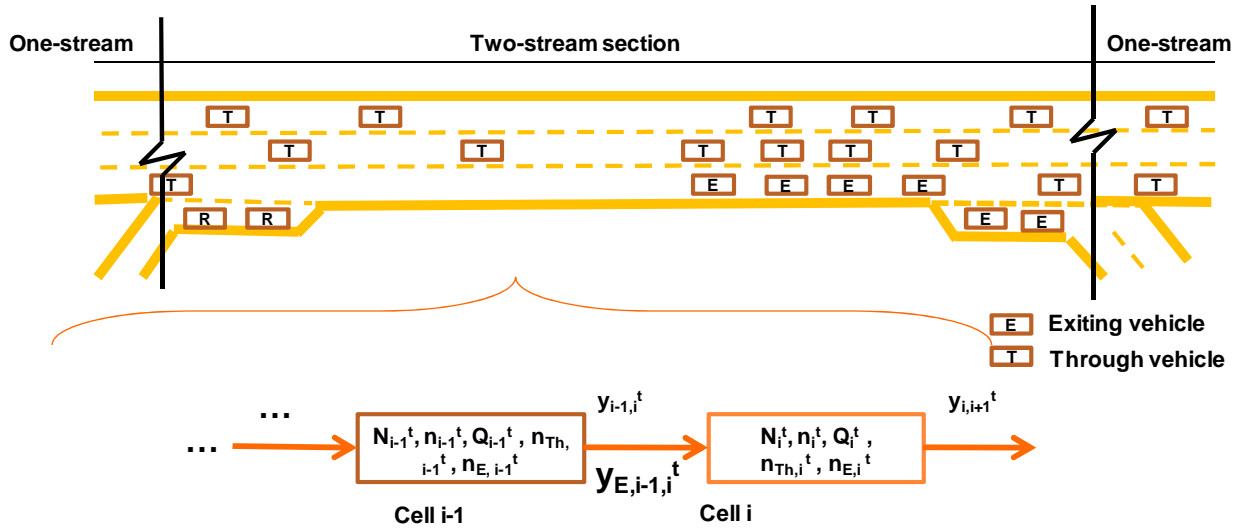
Figure 5-4 Modeling one-stream segment by ordinary cells

$$y_{i,i+1}^t = \left(1 - \gamma_{E,i}^t\right) \times min\{n_i^t, Q_i^t, \delta[N_{i+1}^t - n_{i+1}^t]\} + y_{E,i,i+1}^t \qquad (5\text{-}5)$$

Equation (5-5) determines the total number of vehicles leaving Cell $i+1$ and entering its downstream cell (Cell i). It assumes that those two different traffic streams are well mixed.

$$n_{E,i}^{t+1} = n_{E,i}^t + y_{E,i-1,i}^t - y_{E,i,i+1}^t \qquad (5\text{-}6)$$

Equation (5-3) represents flow conservation law of two-stream cell and Equation (5-6) represents the conservation law of the exit traffic.
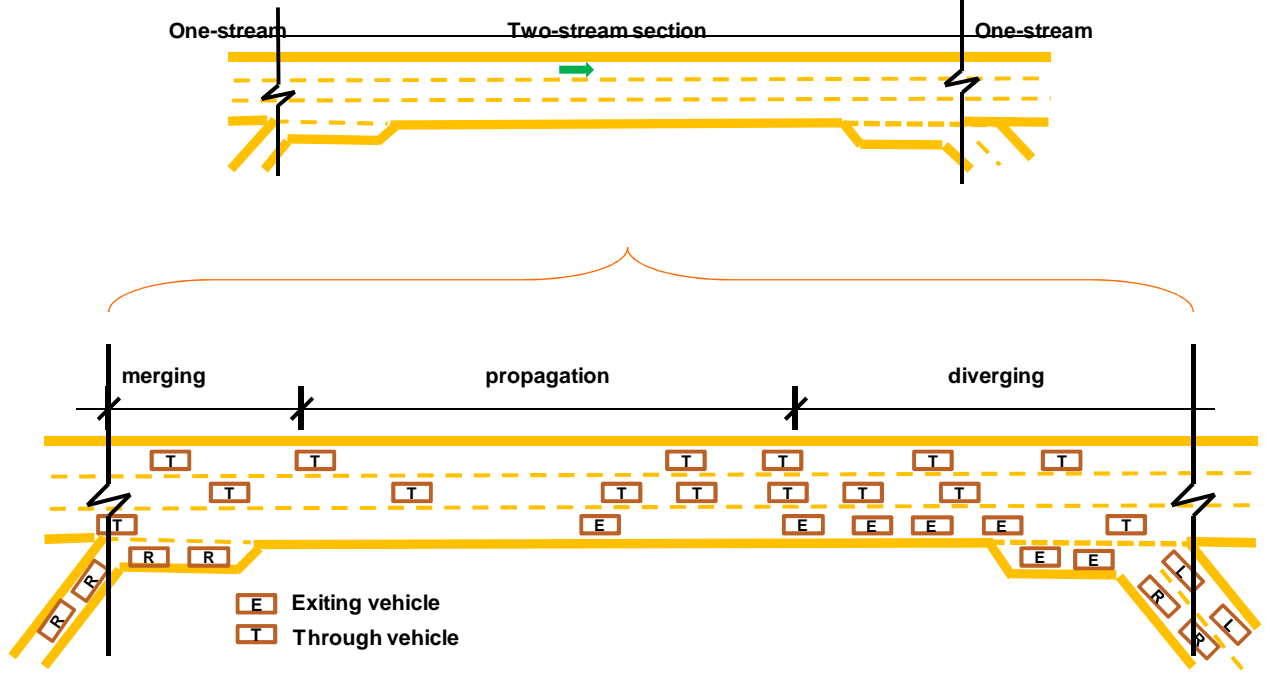
Figure 5-5 Two-stream segment traffic dynamics

As depicted in Figure 5-5, a two-stream freeway segment can be further divided into three zones, namely the merging zone, propagation zone and diverging zone. Among these three zones, the propagation zone can be modeled with the ordinary two-stream cells with sufficient details. The following sections will present the modeling methodologies for merging zone and diverging zone.

Merging zone

As illustrated in Figure 5-6, the merging zone can be represented with a merging cell (Cell C), which has two entry cells, representing the upstream freeway segment and the on-ramp respectively. As the on-ramp traffic should yield to the freeway mainline traffic, the mainline

entry volume (y_{BC}^t) can utilize the capacity first, and the remaining capacity would then be available to on-ramp traffic. Equations (5-7) and (5-8) determine the entry flow from upstream freeway mainline and on-ramp based on the mechanism described above:

$$y_{AC}^t = \min\{n_A^t, Q_A^t, \delta[N_C^t - n_C^t]\} \tag{5-7}$$

$$y_{BC}^t = \min\{n_B^t, Q_B^t, \delta[N_C^t - n_C^t - y_{AC}^t]\} \tag{5-8}$$

$$y_{E,iC}^t = y_{iC}^t \times \xi_E^t, i = A, B \tag{5-9}$$

where ξ_E^t is the pre-determined percentage of vehicles heading to the downstream off-ramp during time interval t.
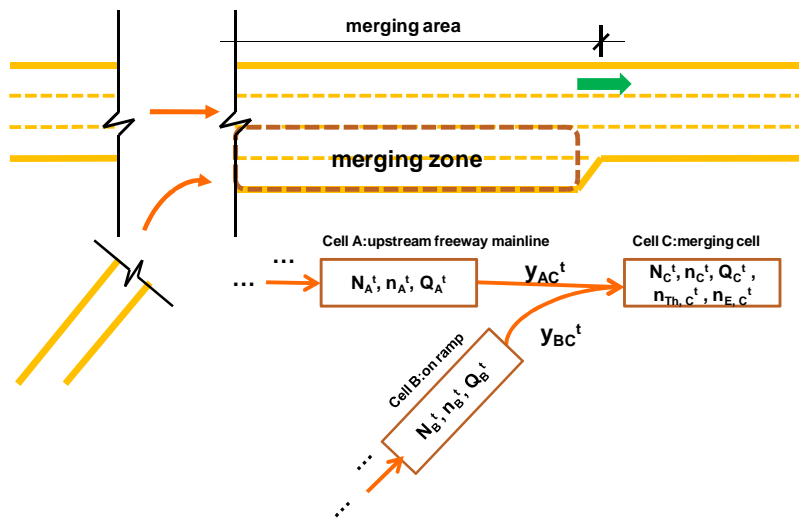
Figure 5-6 On-ramp traffic characteristics

The entry capacity of on-ramp, Q_B^t, is determined by the traffic dynamics in the merging zone shown in Figure 5-6. The length of the acceleration lane and the traffic stream characteristic in the adjacent freeway lane are the two primary factors that may affect Q_B^t. In this study Q_B^t is determined with equation (5-10):

65

$$Q_B^t = (Q_A^t - y_{AC}^t) \times p_C^t \qquad (5\text{-}10)$$

where p_C^t is the lane utilization factor of the right-most lane at time interval t. Equation (5-10) assumes that the on-ramp traffic could use up all the remaining capacity of the right-most freeway lane.

Diverging zone

Under congested conditions, the off-ramp exit queue may spill back to its immediate upstream freeway. There are two major effects on the mainline through traffic when an exit queue spills back, as illustrated in Figure 5-7. Firstly, the exit queue will occupy one or two through lanes, which cannot serve through traffic any more. Secondly, the density of the adjacent mainline lanes will increase, and traffic will slow down due to rubbernecking effect and lateral friction. Therefore, the through capacity of those lanes will diminish.
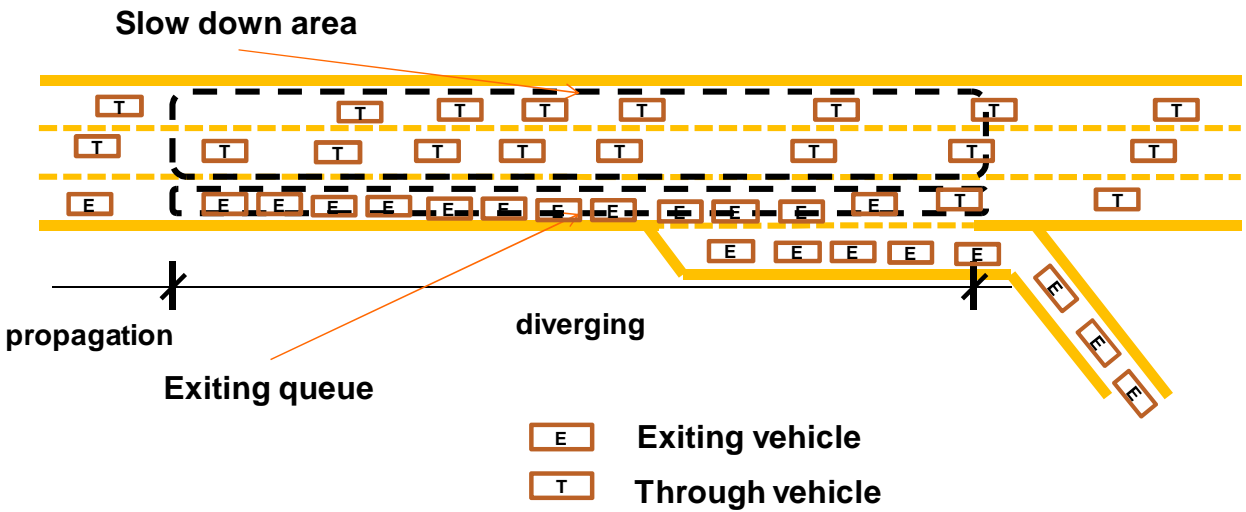
Figure 5-7 The illustration of exit queue effect in diverging zone traffic

To reflect the impact of such complex interactions on the freeway mainline capacity, the diverging zone can be modeled with a diverging cell (named Cell A). There are two downstream cells (Cell B and Cell C), which represent the downstream off-ramp and freeway mainline

segment, respectively (see Figure 5-8). The diverging cell can be further divided into three subareas denoted by Th, TE, and E. All vehicles in subarea Th head to the downstream freeway, and vehicles in subarea E will exit to the adjacent off-ramp. Subarea TE can be shared by both through and exit vehicles. The diverging cell then is divided into two sub-cells (Sub-cell Th and Sub-cell E) to represent the two different streams, as illustrated in Figure 5-8. The buffer capacity of Sub-cell Th, denoted by N_{Th}^t, can be computed as $N_{Th}^t = N_1^t + N_2^t$. That of Sub-cell E, denoted by N_E^t, can be computed as $N_E^t = N_2^t + N_3^t$, where N_1, N_2 and N_3 are the maximum numbers of vehicles that can be present in subzones Th, TE, and E, respectively.
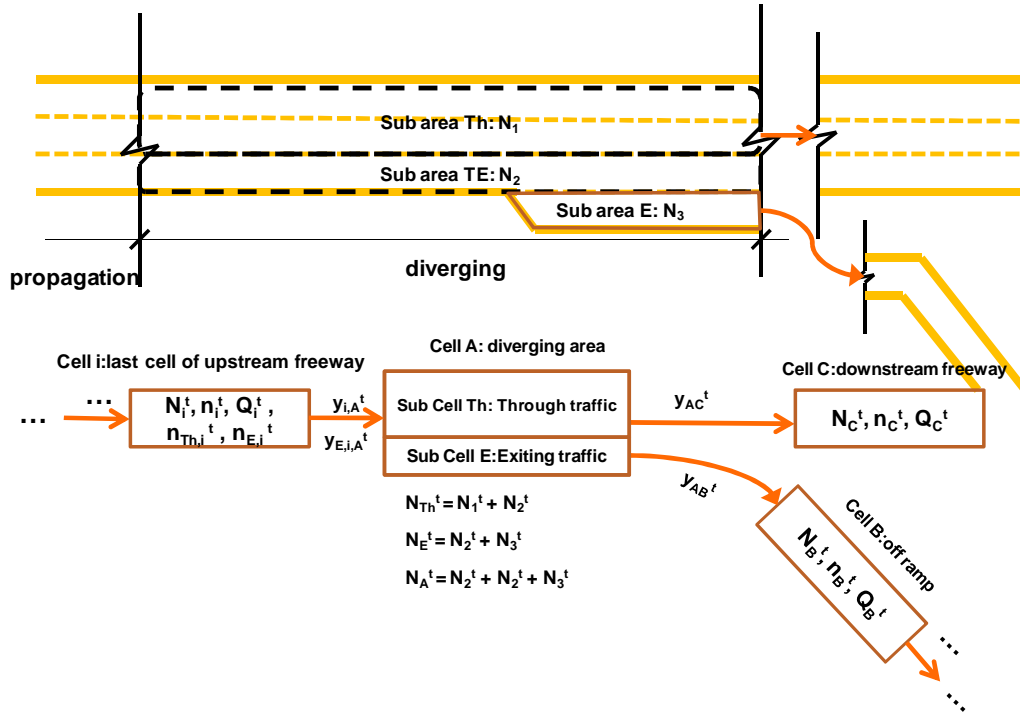
Figure 5-8 Graphical illustration of modeling diverging zone

Equations (5-11)-(5-14) are employed to determine the numbers of through and exit vehicles entering a diverging cell (Cell A) from its upstream cell (Cell *i*):

$$y = min\{n_i^t, Q_i^t, \delta(N_A^t - n_A^t)\} \qquad (5\text{-}11)$$

$$y_{E,i,A}^t = min\{y\gamma_{E,i}^t, N_{E,A}^t - n_{E,A}^t\} \qquad (5\text{-}12)$$

$$y_{i,Th}^t = min\{(1 - \gamma_{E,i}^t) \times y, N_{Th}^t - n_{Th}^t\} \qquad (5\text{-}13)$$

$$y_{iA}^t = y_{i,Th}^t + y_{E,i,A}^t \qquad (5\text{-}14)$$

where y is a temporary variable to simplify description; $y_{i,Th}^t$ stands for the number of vehicles

from cell i to the Sub-cell Th of cell A; $y_{E,i,A}^t$ denotes the exit vehicles which enters Cell A.

Equations (5-12) and (5-13) determine the total number of vehicles from Cell i to the through sub

cell of the diverging cell (Sub-cell Th of Cell A) and the exit sub cell (Sub-cell E of Cell A),

which assumes that the two different traffic streams are well mixed. The exit flows from the sub

cells to their downstream cells follow Equations (5-15) and (5-16). The flow conservation law of

all those cells or sub-cells remains the same.

$$y_{AC}^t = min\{n_{Th}^t, Q_{Th}^t, \delta(N_C^t - n_C^t)\} \qquad (5\text{-}15)$$

$$y_{AB}^t = min\{n_E^t, Q_E^t, \delta(N_B^t - n_B^t)\} \qquad (5\text{-}16)$$

The saturated flow rate of the exit Sub-cell E is clearly equal to the off-ramp saturated

flow rate. However, the computation of the saturated flow rate for the through Sub-cell Th is

more complex. As mentioned earlier, when the exit queue spills back to the freeway mainline,

the traffic on the adjacent lanes will slow down and form a slow speed area due to rubbernecking

and lane-changing effects. This study borrows the rubbernecking concept to model this effect as

follows:

$$Q_{Th}^t = Q_A \times \left(1 - \frac{L_{A,E}}{L_A}\right) \times [1 - \alpha_A \times \frac{n_{E,A}^t}{N_{E,A}^t}] \qquad (5\text{-}17)$$

where $L_{A,E}$ the number of lane is occupied by the exit queue; L_A is total lane number of the

freeway mainline; $Q_A \times (1 - L_{A,E}/L_A)$ is the capacity of the unblocked through lane(s); α_A is the

maximum saturated flow deduction proportion, which is the capacity deduction fraction when

the exit vehicles occupy all the available buffer space of the freeway; $n_{E,A}^t/N_{E,A}^t$ is the proportion

of the exit buffer capacity occupied by the exit vehicles. Equation (5-17) assumes that the through saturated flow rate will be the capacity of remaining though lane(s) with some deduction, and the deduction rate will increase linearly with the length of exit queue with a threshold a_A.

5.3. *An integrated single-interchange control model*

5.3.1 Objective function

Depending on the traffic conditions, we can set the control objective function as maximizing the total system throughput or minimizing the total delay. With the above cell-based formulations, the objective function of maximizing the system throughput can be expressed as follows:

$$Max \ (Throughput = \sum_{t=1}^{T}\sum_{j\in S}\sum_{i\in\Gamma^-(j)} y_{ij}^t) \tag{5-18}$$

where S is the sink cell set, $\Gamma^-(j)$ is the upstream cell set of cell j, and T is total operation time period.

With the CTM method, the length of each cell is set to be the free speed travel distance during a pre-specified time unit, which means that the vehicles in each cell can either stay or move to the downstream cells. If we define the delay as the difference between a vehicle's actual travel time and its free-flow speed travel time over a given travel distance, the delay of a vehicle in a particular cell can be computed as the time interval in which it stay in the same cell. For instance, if some vehicles stay in the same cell over n consecutive unit intervals, then this implies that they all have experienced the delay of n time units. More specifically, one can define the delay for each cell for time interval t as $d_i^t = \left(n_i^{t-1} - \sum_{j\in\Gamma(i)} y_{ij}^t\right)\tau$, where $\Gamma(i)$ is the downstream-cell set of Cell i and τ is the length of one time unit. The alternative objective function of

minimizing the total system delay can be expressed as follows:

$$Min \left[total\ delay = \quad \tau \sum_{t=1}^{T} \sum_{i} \sum_{j \in \Gamma(i)} w_i \left(n_i^t - y_{ij}^t\right)\right] \tag{5-19}$$

where w_i is a weighted coefficient to modify the relative importance of each cell. As τ is a

constant, the objective function of minimizing system delay can be further stated as:

$$Min \quad Z = \sum_{t=1}^{T} \sum_{i} \sum_{j \in \Gamma(i)} w_i \left(n_i^t - y_{ij}^t\right) \tag{5-20}$$

The formulations for signal control and the solution algorithm are identical to those

presented in Chapter 4, except for the inclusion of freeway-related constraints.

5.3.2 Summary

This chapter presents an integrated control method for freeways interchange using the
Cell Transmission concept. The proposed formulations reflect the complex interactions between
ramp queues and mainline vehicles in the merging, propagation, and diverging zones at a typical
freeway interchange. By integrating the arterial signal models with the freeway formulations,
the proposed model can determine the ramp and signal control plan that optimizes the
performance of the entire interchange, including the tradeoff between freeway and arterial delays.

Chapter-6: Numerical Case Study

6.1. Introduction

This chapter presents numerical case studies to demonstrate the performance of the proposed arterial signal optimization model and single interchange control model. The chapter is organized as follows. Section 6.2 presents the numerical case study for the proposed arterial signal optimization model, including a detailed description of the case study site, its traffic demand pattern and design of demand scenarios, the selected GA solver parameters, the method for comparing model performance, and the experiment results with detailed performance compassions. Section 6.3 then covers the numerical case study for the single interchange control model with identical organization.
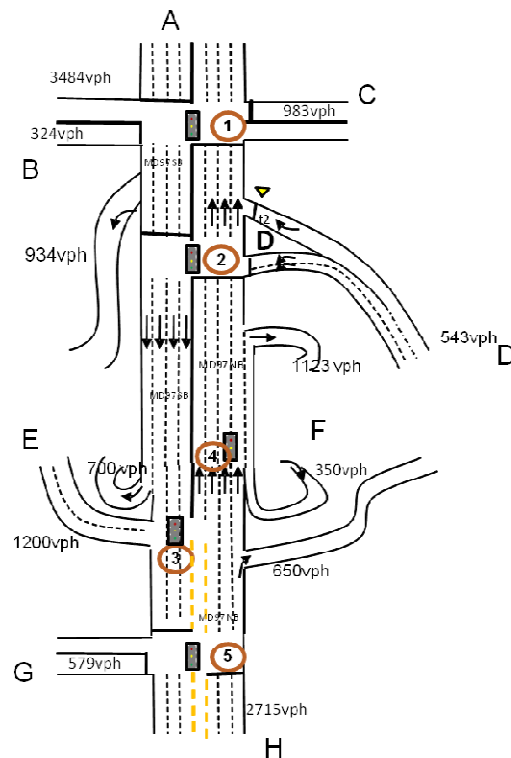
Figure 6-1 Case study site sketch for the arterial signal optimization model

6.2. _Numerical Case study for the proposed arterial signal optimization model_

To evaluate the performance of the proposed model, a segment of Georgia Avenue (MD97) in the Capital Beltway in Silver Spring, Maryland has been selected for the experimental study. As illustrated in Figure 6-1, the target network includes five signalized intersections from Forest Glen Rd (MD192) to Seminary Place. In the case study network, approaches C and D have two left-turn pocket lanes, while B and G have only one left-turn pocket lane.

Using the actual volume (Year 2008) as the base line, the distribution of traffic volume is varied for each approach and three levels of traffic conditions are generated for performance evaluation (see Table 6-1).

Table 6-1 Demands for the case study site (vehicle per hour)

Entrance	Movements	Demand Scenario		
		Low	Medium	High
A	Through	3,044	3,382	3,720
	Right	101	112	123
B	Left	40	44	48
	Through	91	101	111
	Right	161	179	197
C	Left	536	596	656
	Through	306	340	374
	Right	42	47	52
D	Left	284	315	347
	Right	204	227	250
E	Right	1,080	1,200	1,320
F	Right	315	350	385
G	Left	498	553	608
	Right	23	25	28
H	Through	2,444	2,715	2,987
Total	-----	9,167	10,186	11,204

In the above traffic demand scenarios, the medium level volumes are actually from field data collected for the morning peak hour (8:00AM to 9:00AM on 09/10/2008). The low- and

high-level scenarios are 90 percent and 110 percent of the medium level volumes.

The signal plans generated from the proposed model are compared with those generated by Transyt-7F (release 10), which is one of the most advanced programs for both research and practice. Transyt-7F (release 10) offers two optimization algorithms, namely the hill-climb algorithm and the GA algorithm. For a fair comparison, the GA method in Transyt-7F (release 10) has been used to optimize signal timings for the case study. Both GA optimizers run for 200 generations with a population size of 50, a crossover probability of 0.3, and a mutation probability of 0.01.

For comparison, a microscopic simulation, CORSIM, is employed as the performance index provider. For the CORSIM simulation model, O-D (origin-destination) calibration is performed based on the observed demand pattern. At the link level, free flow speed, physical geometry, and saturated flow for each lane-group have been calibrated. The optimization and simulation results are presented in the following sections.

All the simulation runs in the signal optimizers are performed for fifteen minutes as recommended in Highway Capacity Manual 2000 (HCM 2000). All simulation runs follow a three minutes network initialization process.

6.2.1 Resulting signal timings

Table 6-2 summarizes the optimized signal timings (cycle length and offset) for all five signals in the control boundary from both Transyt-7F model and the proposed model. The cycle duration increases with the demand level for both Transyt-7F model and the proposed model. For the same demand level, the proposed model is intended for shorter cycles than the Transyt-7F model. The computation time for the case study is about 20 minutes with one thread in an Intel®

Pentium® D CPU (3.2GHz X 2). For the same condition, the Transyt-7F release needs about 30 minutes.

Table 6-2 Signal timings for the case study site

Demand Scenario	Signal timing		Signal # (second)				
			1	2	3	4	5
Low	Cycle Length	Transyt-7F	99	99	99	99	99
		Proposed Model	78	78	78	39	78
	Offset	Transyt-7F	0	0	12	98	11
		Proposed Model	0	56	77	38	31
Medium		Transyt-7F	120	120	120	120	120
	Cycle Length	Proposed Model	108	54	108	54	108
		Existing conditions	150	150	150	150	150
		Transyt-7F	0	0	22	19	17
	Offset	Proposed Model	0	5	106	51	45
		Existing conditions	80	81	81	81	105
High	Cycle Length	Transyt-7F	182	182	182	182	182
		Proposed Model	123	123	123	123	123
	Offset	Transyt-7F	0	0	172	115	168
		Proposed Model	92	82	0	98	90

6.2.2 Overall system performance comparison

The simulation results from CORSIM for one hour are presented in this section. For each case based on the average of 50 simulation runs, the network-wide total delay, total queue delay, and system throughput are listed in Table 6-3. The results presented in Table 6-3 indicate that the proposed model outperforms Transyt-7F for all three scenarios at the system level. The 95% confidence intervals indicate that the improvements are statistically significant. The delay improvement increases with the congestion level, which implies that the proposed model is especially applicable for optimizing signals under congested conditions. For the medium demand level, the total delay, total queue delay, and throughput from the existing signal timings are also reported. The existing signal timings yield far more total delay and total queue delay than those of the proposed model and Transyt-7F model. The existing timings produce similar total

74

throughputs with Transyt-7F but less than the proposed model. The existing signal timings perform less well than the other two sets of signal timings due to following two reasons: first, the existing signal timings are not optimized for the particular traffic pattern listed in this case study; secondly, there should be some other considerations beyond the system performances which are considered by the existing signal timings but not by the other two models. Therefore, in the remaining comparison, we just compare the proposed model with Transyt-7F.

Table 6-3 Overall model performance comparison

Demand Scenarios		Simulation Results from CORSIM (One hour)					
		Proposed Model	Transyt-7F	Existing Timings	Improvement *	Improvement * (%)	Improvement (95% CI*)
Low	Total Delay (vehicle-hours)	122.34	178.50	--/--	56.16	31%	[28.8,83.5]
	Total Queue Delay (vehicle-hours)*	63.26	105.73	--/--	42.47	40%	[24.4, 60.5]
	Total Throughput (vehicles)	9107	8990	--/--	117	1%	[55, 180]
Medium	Total Delay (vehicle-hours)	174.20	276.01	385.21	101.81	37%	[68.6,135.1]
	Total Queue Delay (vehicle-hours)	100.41	167.10	212.96	66.69	40%	[42.8, 90.6]
	Total Throughput (vehicles)	10047	9870	9081	176	2%	[78, 276]
High	Total Delay (vehicle-hours)	259.11	426.14	--/--	167.03	39%	[135.7,198.3]
	Total Queue Delay (vehicle-hours)	157.99	272.24	--/--	114.25	42%	[95.6, 132.9]
	Total Throughput (vehicles)	10846	10192	--/--	654	6%	[568, 741]

* Delay improvement = Transyt-7F Delay – The Proposed Model Delay
Throughput improvement = The Proposed Model Throughput - Transyt-7F Throughput
Delay Improvement (%) = (Transyt-7F Delay – The Proposed Model Delay) / the Proposed Model Delay × 100%
Throughput Improvement (%) = (The Proposed Model Throughput - Transyt-7F Throughput) / Transyt-7F Throughput × 100%
CI = confidence Interval
Queue delay = Delay calculated by taking vehicles having acceleration rates less than 2 feet per second^2 and speed less than 9 feet per second. If a vehicle's speed is less than 3 feet per second, it will be included every second. Otherwise it will be included every two seconds(ITT Industries 2006).

6.2.3 Delay comparison by intersection and corridor

The total delays for the four intersections, MD 97 SB, and MD 97 NB are presented in

Table 6-4. For the low demand scenario, the proposed model favors the congested intersection (intersection 1), but increases the delay at other intersections. However, the proposed model reduces the total delay experienced by the traffic in MD 97 SB. For the medium and high demand levels, the proposed model can improve the performance of the congested intersections, and the improvement increases with the demand level. For the other intersections, the difference decreases with traffic demand.

Table 6-4 Total delay comparison by intersection (vehicle minutes)

Demand Scenarios		Simulation Results from CORSIM (One hour)				
		Proposed Model	Transyt-7F	Improvement*	Improvement * (%)	Improvement (95% CI)
Low	Intersection 1	2242.41	6112.78	3870.37	63%	[5483.4, 2257.4]
	Intersection 2	588.20	443.24	-144.95	-33%	[-109.1, -180.8]
	Intersection 3	1815.50	1703.63	-111.87	-7%	[-32.1, -191.7]
	Intersection 4	1235.07	1064.19	-170.88	-16%	[-67.5, -274.2]
	MD 97 SB	3951.10	7207.90	3256.80	45%	[1668.9, 4844.7]
	MD 97 NB	1475.10	1637.70	162.60	10%	[101.4, 223.7]
Medium	Intersection 1	3771.26	9376.52	5605.26	60%	[7742.7, 3467.9]
	Intersection 2	537.45	640.41	102.97	16%	[161.3, 44.6]
	Intersection 3	1992.82	2183.95	191.13	9%	[277.0, 105.2]
	Intersection 4	2318.74	1960.66	-358.09	-18%	[-221.3, -494.8]
	MD 97 SB	4261.00	11943.50	7682.50	64%	[5843.2, 9521.8]
	MD 97 NB	2011.00	2010.30	-0.70	-0%	[-84.7, 83.3]
High	Intersection 1	5967.06	16664.37	10697.31	64%	[12008.3, 9386.3]
	Intersection 2	964.92	895.53	-69.39	-8%	[284.9, -423.7]
	Intersection 3	2992.88	2779.33	-213.55	-8%	[57.8, -484.9]
	Intersection 4	2689.92	2590.17	-99.75	-4%	[-8.0, -191.5]
	MD 97 SB	6671.80	17109.80	10438.00	61%	[8845.1, 12031.0]
	MD 97 NB	2854.60	2961.30	106.70	4%	[-73.4, 286.9]

* Delay improvement = Transyt-7F Delay – The Proposed Model Delay
Delay Improvement (%) = (Transyt-7F Delay – The Proposed Model Delay) / the Proposed Model Delay × 100%
CI = confidence Interval

The total delay for MD 97 southbound (SB) and northbound (NB) shows that the proposed model reduces the total delay of MD 97 SB at all three volume levels. For MD 97 NB, the two models provide comparative good performance. The results indicate that the resulting signal timings of the proposed model yield less delay for SB traffic, which has much heavier demand, than for NB traffic.

Table 6-5 Total queue delay comparison by intersection (vehicle minutes)

Demand Scenarios		Simulation Results from CORSIM (One hour)				
		Proposed Model	Transyt-7F	Improvement*	Improvement * (%)	Improvement (95% CI)
Low	Intersection 1	1189.05	3945.16	2756.11	70%	[3859.2, 1653.0]
	Intersection 2	374.54	317.19	-57.35	-18%	[-29.1, -85.6]
	Intersection 3	1288.58	1290.17	1.59	0%	[68.1, -64.9]
	Intersection 4	746.68	617.89	-128.78	-21%	[-33.1, -224.5]
	MD 97 SB	1779.9	4156.2	2376.3	57%	[1288.7, 3463.8]
	MD 97 NB	521.6	747.5	225.9	30%	[187.5, 264.3]
Medium	Intersection 1	2308.33	9376.52	7068.19	75%	[9187.7, 4948.6]
	Intersection 2	288.21	640.41	352.21	55%	[407.2, 297.2]
	Intersection 3	1442.18	2183.95	741.77	34%	[820.4, 663.2]
	Intersection 4	1676.67	1960.66	283.99	14%	[399.6, 168.4]
	MD 97 SB	1708.8	7027.8	5319.1	76%	[4034.5, 6603.6]
	MD 97 NB	875.1	926.3	51.2	6%	[-7.3, 109.8]
High	Intersection 1	3870.18	16664.37	12794.19	77%	[14114.8, 11473.5]
	Intersection 2	594.79	895.53	300.73	34%	[560.6, 40.9]
	Intersection 3	2218.65	2779.33	560.68	20%	[796.8, 324.5]
	Intersection 4	1895.93	2590.17	694.24	27%	[767.1, 621.4]
	MD 97 SB	3049.8	10181.0	7131.2	70%	[6090.4, 8172.0]
	MD 97 NB	1379.7	1464.1	84.4	6%	[-54.1, 223.0]

* Delay improvement = Transyt-7F Queue Delay – The Proposed Model Queue Delay
Delay Improvement (%) = (Transyt-7F Queue Delay – The Proposed Model Queue Delay) / the Proposed Model Queue Delay × 100%
CI = confidence Interval

Table 6-5 summarizes the total queue delay for each intersection, MD 97 SB, and MD 97

NB. It indicates that the proposed model reduces the total queue delay for the most congested intersection (Intersection 1) at all three demand levels. For other intersections, the proposed model's performance improves with the demand level. For the congested corridor (MD97 SB), the proposed model can produce less queue delay than Transyt-7F. For the opposite direction (MD97 NB), the total queue delays from both models are comparative.

For all three demand levels, the proposed model provides better performance than Transyt-7F with respect to the total system delay and total system throughput. The improvement seems to increase with the demand level. That is the advantage of tracking the movement blockage since the probability of incurring a blockage increases with demand. By tackling the traffic dynamics in a more accurate way, the proposed model reduces the total delay experienced by the traffic on MD 97 SB, and MD 97 NB. The results demonstrate that the proposed model is promising for oversaturated traffic conditions.

6.2.4 Conclusions

Extensive simulation experiments for field segment of four congested intersections have demonstrated that both the total delay and throughput resulting from the proposed model are better than those with Transyt-7F under a wide range of traffic conditions, especially at high traffic volumes. Hence, the proposed model is ready for use in practice, as illustrated by the case study, especially under oversaturated conditions.

6.3. _Case study for the proposed single interchange control model_
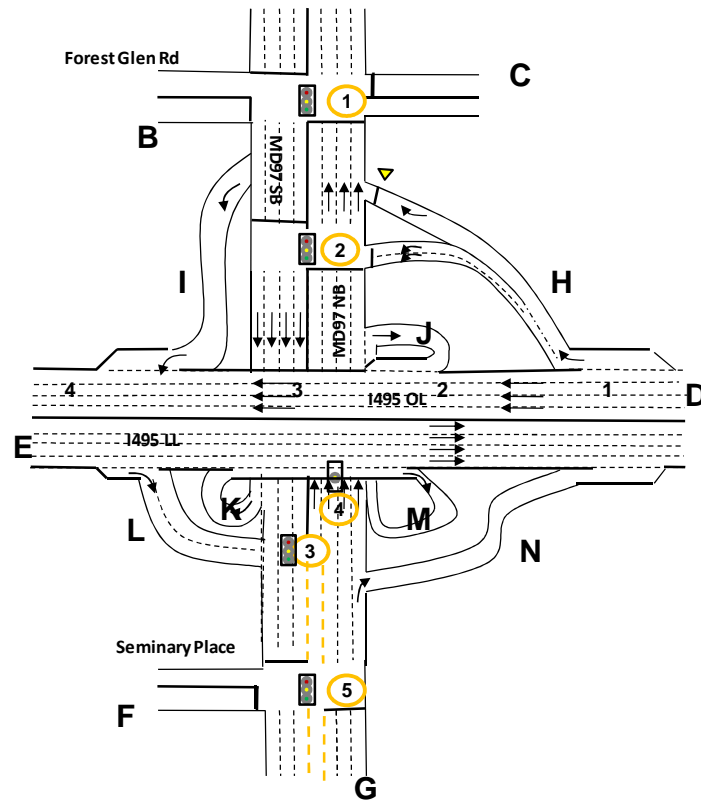
Figure 6-2 Case study site sketch for the single interchange model

To evaluate the performance of the proposed model, this study has selected the Capital Beltway (I-495) / Georgia Avenue (MD97) interchange in Silver Spring, Maryland. As shown in Figure 6-2, the target site includes five signalized intersections from Forest Glen Rd (MD192) to Seminary Place, among which, signals 2, 3, and 4 are the interchange signals. The major road segments in this case study site include I-495 Outer Loop (I-495 OL), I-495 Inner Loop (I-495 IL), MD 97 Southbound (MD 97 SB), and MD 97 Northbound (MD97 NB).

The actual entrance volumes, which are presented in Table 6-6, are based on detector data. Based on the same entry traffic, three scenarios are defined with different exit volumes from off-ramp H. These scenarios have low, medium, and high exit volumes of 1400vph, 1600vph, and

1800vph, respectively. To simplify description, the following sections refer to them as the low, medium, and high scenario. Table 6-6 also presents the entrance volumes and the turn volumes for each diverging point under each scenario.

Table 6-6 Basic demand for the case study of interchange control model

Entrance		Movements	Exit (Ramp-H) Volume Scenario (vehicles per hour)		
			Low	Medium	High
Entrance	A	Through	3,382	3,382	3,382
		Right	112	112	112
	B	Left	44	44	44
		Through	101	101	101
		Right	179	179	179
	C	Left	596	596	596
		Through	340	340	340
		Right	47	47	47
	D	Through	7025	7025	7025
	E	Through	6879	6879	6879
	F	Left	553	553	553
		Right	25	25	25
	G	Through	2715	2,715	2,715
	Total	-----	21998	21998	21998
Ramp	H	Right	**586**	**670**	**754**
		Left	**814**	**930**	**1046**
		Total	**1400**	**1600**	**1800**
	I	Enter	825	825	825
	J	Enter	1402	1402	1402
	K	Enter	829	829	829
	L	Exit	1179	1179	1179
	M	Exit	298	299	300
	N	Enter	654	654	654

The signal plans generated from the proposed model are compared with those from Transyt-7F, which is one of the most advanced programs for both research and practice, and often used as a reference method to test improvements of signal optimization models (Papageorgiou et al. 2003). Transyt-7F (release 10) offers two optimization algorithms, namely the hill-climb algorithm and the GA algorithm. For a fair comparison, this study uses the GA

method of Transyt-7F (release 10) to optimize signal timings with same parameters. The GA

optimizers in both Transyt-7F and the proposed model run for 200 generations with a population

size of 50, a crossover probability of 0.3, and a mutation probability of 0.01. The simulation

period is set at 15 minutes for both optimizers, which is recommended by Highway Capacity

Manual 2000 (HCM 2000). A 3-minute network initialization process is used for all programs.

Table 6-7 Signal timings for the case study site

Demand Scenario	Signal timing		Signal # (second)				
			1	2	3	4	5
Low	Cycle Length	Transyt-7F	90	90	90	90	90
		Proposed Model	174	87	87	174	174
	Offset	Transyt-7F	0	53	46	51	56
		Proposed Model	0	68	11	6	47
Medium	Cycle Length	Transyt-7F	160	160	160	160	160
		Proposed Model	70	35	70	35	35
	Offset	Transyt-7F	0	80	79	81	65
		Proposed Model	0	18	22	27	29
High	Cycle Length	Transyt-7F	135	135	135	135	135
		Proposed Model	82	41	82	41	82
	Offset	Transyt-7F	0	50	34	47	25
		Proposed Model	0	32	3	5	40

Table 6-7 summarizes the optimized signal timings (Cycle length and offset) for the five

signals in the control boundary from both Transyt-7F and the proposed model. For the same

demand level, the proposed model intends to use shorter cycles than Transyt-7F. The

performance comparisons resulting from CORSIM are presented below.

6.3.1 Experimental Results

Table 6-8 presents the average network-wide MOEs (Measurement of Effectiveness)

during one hour simulation from 50 CORSIM runs with the signal timings produced by the

proposed model and Transyt-7F. These results clearly indicate that the proposed interchange

model produces less total delay than Transyt-7F for all three scenarios, and the improvements are

statistically significant as indicated by the 95% confidence intervals. The results also indicate a trend whereby delay improvement increases with off-ramp volume.

Table 6-8 Overall model performance comparison

Demand Scenarios	MOEs	Simulation Results from CORSIM (One hour)				
		ICIC*	TY7F*	Improvement*	Improvement * (%)	Improvement (95% CI*)
Low	Total Delay (vehicle-hour)	545.6	617.1	71.4	11.6%	[66.0, 76.8]
	Total Throughput (vehicle)	20591	20473	117	0.6%	[91, 144]
Medium	Total Delay (vehicle-hour)	464.7	931.5	466.8	50.1%	[460.7, 472.9]
	Total Throughput (vehicle)	20922	18098	2824	15.6%	[2777, 2870]
High	Total Delay (vehicle-hour)	427.8	943.2	515.5	54.6%	[504.2, 526.8]
	Total Throughput (vehicle)	20939	19057	1882	9.9%	[1801, 1964]

* ICIC: the proposed Interchange Integrated Control Model
  TY7F: Transyt-7F
  Delay improvement = Transyt-7F Delay – The Proposed Model Delay
  Throughput improvement = The Proposed Model Throughput - Transyt-7F Throughput
  Delay Improvement (%) = (Transyt-7F Delay – The Proposed Model Delay) / the Proposed Model Delay × 100%
  Throughput Improvement (%) = (The Proposed Model Throughput - Transyt-7F Throughput) / Transyt-7F Throughput × 100%

The proposed model produces almost the same system throughput as Transyt-7F under the low off-ramp volume scenario, as illustrated in Table 6-8. However, as the exit volume from Ramp-H increases in the medium and high scenarios, Transyt-7F model produces less system throughput and more total delay. The proposed model produces slightly more system throughput and much less total system delay for those two scenarios.

Table 6-9 Total delay comparison by roadway segment (vehicle minutes)

Demand Scenarios			Simulation Results from CORSIM (One hour)				
			ICIC	TY7F	Improvement*	Improvement (%)	Improvement (95% CI)
Low	Freeway	I-495IL	5219.0	5124.3	-94.7	-1.8%	[-220.0, 30.7]
		I-495OL	10807.5	10654.7	-152.8	-1.4%	[-313.4, 7.8]
		Total	16312.2	16091.9	-220.3	-1.4%	[4330.8, 4908.0]

82

	Arterial	MD97NB	2795.9	2866.2	70.4	2.5%	[49.9, 90.9]
		MD97SB	7439.4	14811.7	7372.2	49.8%	[7141.3, 7603.2]
		Total	16426.4	20931.6	4505.3	21.5%	[4232.0, 4778.6]
Medium	Freeway	I-495IL	5294.9	4390.0	-904.9	-20.6%	[-972.5, -837.2]
		I-495OL	9390.7	11195.1	1804.4	16.1%	[1465.0, 2143.8]
		Total	14993.0	15997.9	1004.9	6.3%	[657.5, 1352.3]
	Arterial	MD97NB	2537.7	4186.6	1649.0	39.4%	[1623.7, 1674.2]
		MD97SB	6162.3	29802.4	23640.1	79.3%	[23421.6, 23858.6]
		Total	12891.2	39892.8	27001.6	67.7%	[26732.3, 27270.9]
High	Freeway	I-495IL	5326.3	10220.3	4894.0	47.9%	[4573.6, 5214.4]
		I-495OL	6842.6	9665.4	2822.7	29.2%	[2505.1, 3140.4]
		Total	12500.3	20552.8	8052.5	39.2%	[7595.5, 8509.6]
	Arterial	MD97NB	2435.6	5520.5	3084.9	55.9%	[3065.3, 3104.5]
		MD97SB	6318.1	22999.8	16681.7	72.5%	[16291.2, 17072.3]
		Total	13165.7	36041.1	22875.4	63.5%	[22453.9, 23296.9]

* Delay improvement = Transyt-7F Delay – The Proposed Model Delay

Delay Improvement (%) = (Transyt-7F Delay – The Proposed Model Delay) / the Proposed Model Delay × 100%

CI = confidence Interval

Table 6-9 presents total delays on both freeways and arterials for all three scenarios. For the low scenario, the proposed model yields slightly more freeway delay (16312 vs. 16091 vehicle-minutes) but far less arterial delay (16426 vs. 20931 vehicle-minutes) than Transyt-7F. For both medium and high scenarios, the proposed model produces far less freeway and arterial delays. The improvements in freeway delay increase with the exit volume from Ramp-H (-220.3, 1004.9, and 8052.5 vehicle-minutes for low, medium, and high scenario). The improvements in arterial delay increase from 4505.3 vehicle-minutes for the low scenario to 27001.6 vehicle-minutes for the medium scenario, and then drop to 22875.4 vehicle-minutes for the high scenario. It indicates that Transyt-7F does not achieve an optimal solution for the medium scenario with the specified GA iterations.

Freeway Delays (vehicle-minutes)
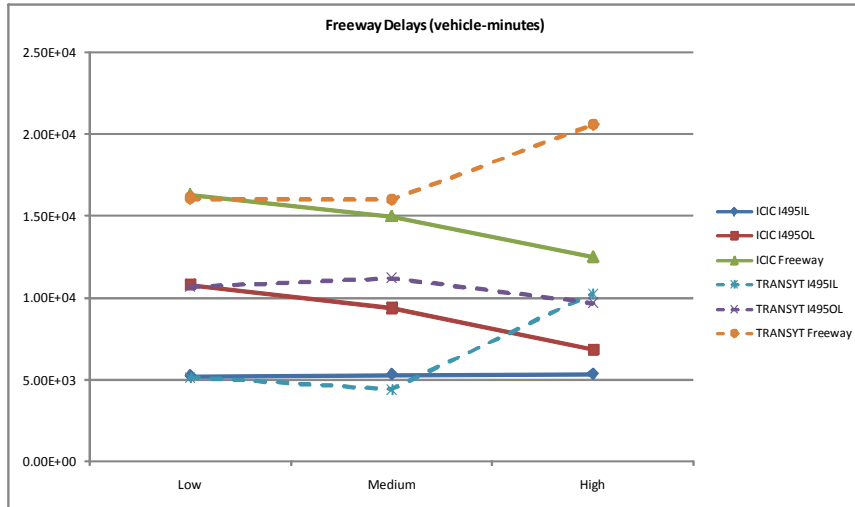
(a) Freeway delay comparison

(b) Arterial delay comparison

Arterial Delays (vehicle-minutes)
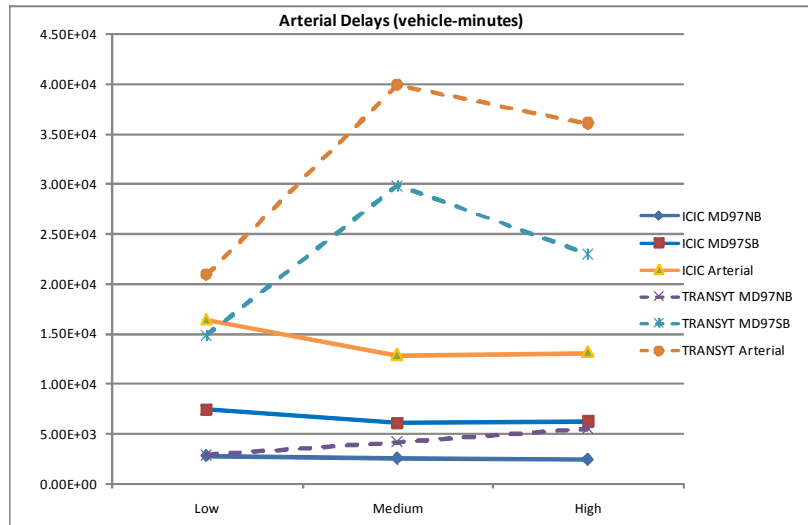
Figure 6-3 Freeway and arterial delay comparison

Figure 6-3(a) illustrates the relation between the freeway delay and the exit volume from Ramp-H for both models. With the proposed model, the delay on I-495 IL is relatively stable and that of I-495 OL decreases with the exit volume from Ramp-H. The proposed model optimizes the traffic signal timings in the context of both freeways and arterials delay, and aims to prevent an off-ramp queue from spilling back to its upstream freeway mainline. The constant I-495 IL delay indicates that its traffic conditions are not impacted by the Ramp-H volume using the control strategies from the proposed model. The decrease of I-495 OL delay reveals that the traffic conditions on I-495 OL improve when Ramp-H volume increase, which can be interpreted

84

that less traffic on I-495 OL improves its traffic conditions. These trends indicate that the proposed model successfully prevents off-ramp queues from spilling back to their upstream freeway mainline. Otherwise, the freeway delays should increase to some degree.

With the Transyt-7F model, the situation is mixed since Transyt-7F does not take the off-ramp spillback into account when optimizing signal timings. The I-495 OL delay does show some trend to decreasing trend as Ramp-H volume increases, as illustrated in Figure 6-3(a). However, the I-495 IL delay jumps from 4390.0 vehicle-minutes to 10220.3 vehicle-minutes, which is a clear indication of queue spillback from Ramp-H. This explains why the Transyt-7F model yields about 40 percent more freeway delays than the proposed model.

Figure 6-3(b) indicates that the proposed model yields less arterial delays than Transyt-7F for all three scenarios. The total arterial delays are 16426.4 vs. 20931.6 (the proposed model vs. Transyt-7F model), 12891.2 vs. 39892.8, and 13165.7 vs. 36041.1 vehicle-minutes for low, medium, and high scenario, respectively. The arterial delay improvement reaches its peak under the medium scenario. With the proposed model, the delay of MD97 NB stays stable and that of MD97 SB decreases slightly, although more traffic enters the arterial system when the exit volume of Ramp-H increases. It is reasonable to suspect that Transyt-7F does not achieve an optimal solution for the medium scenario with the given GA parameters since the arterial delays for the medium scenario are unreasonably above those of the high scenario with Transyt-7F timings.

Table 6-10 compares queue delays of the five intersections, MD97 SB, MD97 NB, and arterials for all three demand levels. These results clearly indicate that the proposed model produces less arterial queue delays for all three scenarios, which are confirmed to be significant by the confidence intervals. The queue delays determined by the proposed model for MD97 SB,

which is the most congested corridor, are significantly below those from Transyt-7F. For MD97 NB, the proposed model yields significantly less queue delay for medium and high volume levels, but slightly more for the low volume level. The proposed model reduces the queue delay of MD97 SB at the cost of MD97 NB.

Table 6-10 Queue delay comparison by intersection (vehicle minutes)

Demand Scenarios		Queue delay* results from CORSIM (One hour)				
		ICIC	TY7F	Improvement*	Improvement (%)	Improvement (95% CI)
Low	Intersection 1	3669.8	8769.3	5099.5	58.2%	[4928.8, 5270.3]
	Intersection 2	1383.1	1154.9	-228.2	-19.8%	[-281.8, -174.6]
	Intersection 3	1647.2	1476.1	-171.1	-11.6%	[-1234.6, -1218.8]
	Intersection 4	515.1	454.1	-61.0	-13.4%	[-219.3, -122.8]
	Intersection 5	2102.7	876.0	-1226.7	-140.0%	[-68.3, -53.6]
	MD97NB	1469.1	1365.5	-103.6	-7.6%	[-116.9, -90.3]
	MD97SB	3772.3	9155.3	5383.0	58.8%	[5216.0, 5550.0]
	Arterials	10465.0	13081.0	2616.0	20.0%	[2415.8, 2816.2]
Medium	Intersection 1	2359.5	20175.3	17815.8	88.3%	[17696.5, 17935.1]
	Intersection 2	919.6	4697.3	3777.7	80.4%	[3741.9, 3813.4]
	Intersection 3	1635.1	1918.4	283.3	14.8%	[-492.8, -382.7]
	Intersection 4	432.8	405.0	-27.7	-6.9%	[250.8, 315.9]
	Intersection 5	1333.9	896.1	-437.8	-48.9%	[-34.9, -20.6]
	MD97NB	1160.2	2441.3	1281.1	52.5%	[1264.3, 1297.9]
	MD97SB	2844.9	25951.7	23106.8	89.0%	[22996.0, 23217.7]
	Arterials	7336.0	33441.0	26105.0	78.1%	[25942.3, 26267.6]
High	Intersection 1	2299.8	9850.9	7551.1	76.7%	[7277.5, 7824.6]
	Intersection 2	1018.9	2933.9	1914.9	65.3%	[1840.0, 1989.8]
	Intersection 3	1569.4	2805.8	1236.4	44.1%	[4895.2, 4953.7]
	Intersection 4	279.0	445.8	166.8	37.4%	[1196.0, 1276.8]
	Intersection 5	1615.3	6539.7	4924.4	75.3%	[160.1, 173.5]

MD97NB	1017.3	3261.9	2244.6	68.8%	[2231.8, 2257.4]
MD97SB	2911.7	15986.0	13074.3	81.8%	[12806.1, 13342.5]
Arterials	7457.3	25781.1	18323.8	71.1%	[18022.1, 18625.5]

* Queue Delay: Delay calculated by taking vehicles having acceleration rates less than 2 feet per second2 and speed less than 9 feet per second. If a vehicle's speed is less than 3 feet per second, it will be included every second. Otherwise it will be included every two seconds.
Delay improvement = Transyt-7F Queue Delay – The Proposed Model Queue Delay
Delay Improvement (%) = (Transyt-7F Queue Delay – The Proposed Model Queue Delay) / the Proposed Model Queue Delay × 100%
CI = confidence Interval

The proposed model reduces queue delay for the most congested intersection (Intersection 1) significantly at all three volume levels, and reduces queue delays of the other four arterial intersections at high volume, but increases them at low volume. The proposed model intends to favor the most congested intersection (Intersection 1).

The comparative analysis of CORSIM simulation results demonstrates that the proposed model outperforms Transyt-7F when optimizing the traffic signal timings around a congested interchange. The proposed model improves total delays, system throughputs, and queue delays for both freeways and arterials for all the three scenarios. The detailed analysis also reveals that the proposed model successfully prevents off-ramp queue spillback for the medium and high scenarios.

6.3.2 CONCLUSIONS

To demonstrate the performance of the proposed integrated control model, extensive simulation experiments are conducted for the Capital Beltway / Georgia Avenue (MD97) interchange in Silver Spring, MD. The results demonstrate that the proposed model produces less total delay and more throughput than Transyt-7F under a wide range of traffic conditions, especially for high off-ramp volumes. The proposed model successfully prevents off-ramp queues from spilling back to their upstream freeway mainline for the medium and high scenarios. It can be concluded that by optimizing the adjacent signals of a congested interchange, the

overall system performance as well as that of individual freeway and arterial could be improved. These results confirm that it is highly desirable to jointly control the freeways and arterials around a congested interchange in order to improve system performance.

6.4. *Conclusions and Limitations*

This chapter shows that the proposed models can capture the link- and lane-blockage on congested arterials, as well as the on-ramp and off-ramp spillback around an interchange. Ideally, we should compare the field traffic data against those from CTM simulation. However, these kinds of validation require massive efforts and resources in collecting field data, which are unavailable for this dissertation. In this dissertation, we reached our conclusion that the proposed models outperform the Transyt-7F model in providing better traffic dynamics by comparing the resulting signal timings from the proposed models with those from Transyt-7F. However, since all those parameters and comparison MOEs are from CORSIM instead of field traffic data, we should restate our conclusion to state that the proposed models are better than Transy-7F model in representing the traffic model of CORSIM.

Although this chapter can only show the advantages of the proposed models in replicating CORSIM simulation results, the results are still very important and valuable for the following reason. In this dissertation, we use all the parameters from CORSIM simulation models for both our proposed models and Transyt-7F. We also compare the resulting signal timings with CORSIM simulation and find that the proposed models outperform Transyt-7F with respect to CORSIM MOEs. Therefore, we can confidently conclude that the proposed models represent the traffic scenarios defined by these parameters and the CORSIM driver behavior model since the signal timings from the proposed models provide better CORSIM MOEs than those from Transyt-7F. CORSIM is among the most popular microscopic simulation packages

for both academia and practice, and has been employed to compare alternatives in many practical projects. Therefore, the results in CORSIM can reflect the traffic conditions in the field with proper calibration. We can reasonably believe that with another set of CORSIM parameters, the proposed models can also perform better than Transyt-7F.

Chapter-7:   Multi-interchange control model

7.1. _Introduction_

This chapter focuses on the development of a multi-interchange integrated control model for freeways and their neighboring arterials under freeway congestion. To develop such a model, it is essential to formulate more realistically the exit queue impact on the mainline through traffic.

To realistically represent the traffic dynamics with freeway mainline spillback, the impact of an exit queue to the adjacent through traffic should be considered. This chapter aims to develop a capacity reduction model for freeway through traffic with an exit queue. The proposed model considers the lateral effect by introducing a friction term to the car-following model proposed by Gazis et al. (1959).

The remaining parts of this chapter are organized as follows: Section 7.2 presents the proposed car-following model with lateral effect and Section 7.3 derives the capacity reduction model for freeway through traffic with an exit queue. Section 7.4 briefly describes the multi-interchange control model. A detailed numerical case study is included in Section 7.5 and conclusions are presented in Section 7.6.

7.2. _A car following model considering lateral effect_

The effect of the neighboring lane traffic on the drivers' behavior has long been omitted since, under light traffic, this effect can be omitted. However, the behavior of those drivers traveling with a long exit queue in the neighboring lane can be affected by the exit queue in the following two aspects. Firstly, the lane-changing from the exit queue to the through lanes makes the drivers in the through lane more cautious, thus reducing the moving lane traffic velocity;

secondly, drivers feel unsafe driving at high speed with a long vehicle queue due to some complex psychological mechanism. Those two factors could reduce the capacity of the through lanes.

To model this effect, the concept of viscosity is borrowed from fluid dynamics. The viscosity is the fluid resistance to shear, which is caused by intermolecular friction exerted when layers of fluids attempt to slide by one another. By Newton's Law of Friction, the shearing stress between the layers of non turbulent fluid moving in straight parallel lines can be defined with the following equation:

$$\tau = v \frac{\partial u}{\partial y} \qquad (7\text{-}1)$$

where τ is the shearing stress; v is the dynamic viscosity, which is a fluid characteristic coefficient; $\partial u/\partial y$ is fluid flow velocity gradient in the direction perpendicular to the flow layers.

Inspired by this idea, the car follow-the-lead model can be modified to account for the psychological friction induced by the long queue in the neighboring lane. The resulting car following model can be further integrated into a flow-density relation, from which we can further derive a capacity model considering the lateral queue effect:

$$\ddot{x}_{n+1}(t) = \frac{\lambda(\dot{x}_n - \dot{x}_{n+1})}{x_n - x_{n+1}} \qquad (7\text{-}2)$$

Equation (7-2) represents the car-follow-the-lead model derived by Gazis et al. (1959) from the speed-density relationship of Greenberg (1959), which has been verified with Lincoln Tunnel data. Equation (7-2) assumes that the acceleration of the following car is proportional to the relative velocity of the two cars. In Equation (7-2), $\dot{x}_n - \dot{x}_{n+1}$ is the relative velocity of the leading car and the following car; $\lambda/(x_n - x_{n+1})$ represents the sensitivity which is inversely

proportional to the space headway.

Considering the driver's perception time, τ, the car-follow-the-lead model can be represented by the following equations:

$$\ddot{x}_{n+1} = \left[\frac{\lambda(\dot{x}_n - \dot{x}_{n+1})}{x_n - x_{n+1}} \right]_{\tau-1} \tag{7-3}$$

$$\ddot{x}_{n+1} = \left[\frac{\lambda(\dot{x}_n - \dot{x}_{n+1})}{x_n - x_{n+1}} \right]_{\tau-1} - \alpha(\dot{x}_{n+1} - \dot{x}_{queue}) \tag{7-4}$$

Equation (7-3) indicates that the following car will change its acceleration in response to the stimuli from the leading car after receiving the stimuli. Equation (7-4) is employed to capture the lateral effect from the queue of the neighbor lane, which follows Newton's Law of Friction. In Equation (7-4), the second term, $-\alpha(\dot{x}_{n+1} - \dot{x}_{queue})$, represents the lateral effect. In the second term, the minus sign indicates that the lateral effect is a resistance; $(\dot{x}_{n+1} - \dot{x}_{queue})$ is the relative speed of the vehicle and the adjacent queue; α is a coefficient reflecting the vehicle's physical parameters and the road conditions.

7.3. *The capacity model for freeway main line with exit queue*

In this section, a capacity model for a freeway main line with exit queue is proposed based on the car-following model presented above. If T is the equilibrium time, the following relation exists, i.e., the speed change for the vehicle is the integral of its acceleration during the time period until equilibrium is reached:

$$\int_0^T \ddot{x}_{n+1} d\tau = v - v_0 \tag{7-5}$$

On the other hand, the integral of the acceleration can be obtained by integrating the

93

right-hand side of Equation (7-4) as follows:

$$\int_0^T \ddot{x}_{n+1} d\tau = \int_0^T \left[\frac{\lambda(\dot{x}_n - \dot{x}_{n+1})}{x_n - x_{n+1}} \right]_{\tau-1} d\tau - \alpha \int_0^T \dot{x}_{n+1} d\tau$$

$$= \lambda \ln(x_n - x_{n+1})|_{\tau=0}^{\tau=T} - \alpha L = \lambda \left[\ln \frac{x_n(T) - x_{(n+1)}(T)}{x_n(0) - x_{n+1}(0)} \right] - \alpha L$$

(7-6)

In Equation (7-6), $L = \int_0^T \dot{x}_{n+1} d\tau$, and represents the distance that the mainline vehicle travels from the start of the queue to the equilibrium position. By definition, the traffic density can be computed as the reciprocal of the space headway. Equation (7-7) represents this relationship. By substituting Equation (7-7) to Equation (7-6), the integral of its acceleration can be further represented by Equation (7-8):

$$k(T) = \frac{1}{x_n(T) - x_{n+1}(T)}$$

(7-7)

$$\int_0^T \ddot{x}_{n+1} d\tau = \lambda \left(\ln \frac{1}{k(T)} - \ln \frac{1}{k(0)} \right) - \alpha L$$

(7-8)

Substituting Equation (7-5) into Equation (7-8), we obtain

$$v - v_0 = \lambda \ln \frac{k(0)}{k(T)} - \alpha L$$

(7-9)

A special steady traffic state is that at jam density (k_j), the corresponding traffic velocity should be zero $(v_0 = 0 \ when \ k(0) = k_j)$. Substituting this into Equation (7-12):

$$v = \lambda \ln \frac{k_j}{k(T)} - \alpha L$$

(7-10)

From the basic flow-density-velocity relationship, we can derive the following equation:

$$q = kv = k\lambda \ln \frac{k_j}{k(T)} - \alpha L k \qquad (7\text{-}11)$$

The first and second derivatives of q in terms of k can be expressed as:

$$\frac{dq}{dk} = \lambda \ln \frac{k_j}{k} - \lambda - \alpha L \qquad (7\text{-}12)$$

$$\frac{d^2q}{dk^2} = -\frac{\lambda}{k_j} < 0 \qquad (7\text{-}13)$$

Since q is a strict convex function of k ($\frac{d^2q}{dk^2} < 0$), it reaches its maximum at the extreme point

($\frac{dq}{dk} = 0$), which can be viewed as the road capacity. Let $\frac{dq}{dk} = 0$,

$$\frac{dq}{dk} = \lambda \ln \frac{k_j}{k^*} - \lambda - \alpha L = 0 \qquad (7\text{-}14)$$

Solving Equation (7-14),

$$k^* = k_j e^{-\left(1+\frac{\alpha L}{\lambda}\right)} \qquad (7\text{-}15)$$

We can find the corresponding flow (capacity) as:

$$q_L^* = k^* v^* = \lambda k_j e^{-\left(1+\frac{\alpha L}{\lambda}\right)} \qquad (7\text{-}16)$$

If there is no queue in the neighboring lane (L=0), the corresponding capacity is:

$$q_0^* = k^* v^* = \lambda k_j e^{-1} \qquad (7\text{-}17)$$

where q_0^* is the capacity with no neighboring queue. The relation of though-lane capacity with or

without neighboring exit queue can be represented as:

$$q_L^* = k^* v^* = q_0^* e^{-\frac{\alpha L}{\lambda}}$$ (7-18)

The first order Taylor approximation is

$$q_L^* = q_0^* e^{-\frac{\alpha L}{\lambda}} \approx q_0^* (1 - \frac{\alpha}{\lambda} L)$$ (7-19)

This is the formulation employed in Chapter 5, which is the first order Taylor approximation and suitable for relatively short queues.

7.4. *Multiple interchange control model*

The capacity reduction model presented in Equation (7-18) for a freeway mainline enables the single interchange control model from Chapter 5 to track the spillback to an upstream interchange caused by the exit queue, which extends the ability of the single interchange control model to optimize signal timings for multiple adjacent interchanges.

7.5. *Numerical case study*

7.5.1 Case study site description

To evaluate the performance of the proposed multi-interchange model, a segment of the Capital Beltway in Silver Spring, Maryland has been selected for the experimental study, which includes four interchanges with four major, namely Georgia Avenue (MD97), Columbia Pike (US 29), University Blvd (MD 193), and New Hampshire Ave (MD 650), as illustrated in Figure 7-1. Among these four interchanges, the distance from US 29 to MD 193 is about 0.5 miles, which indicates easy spillback between them.
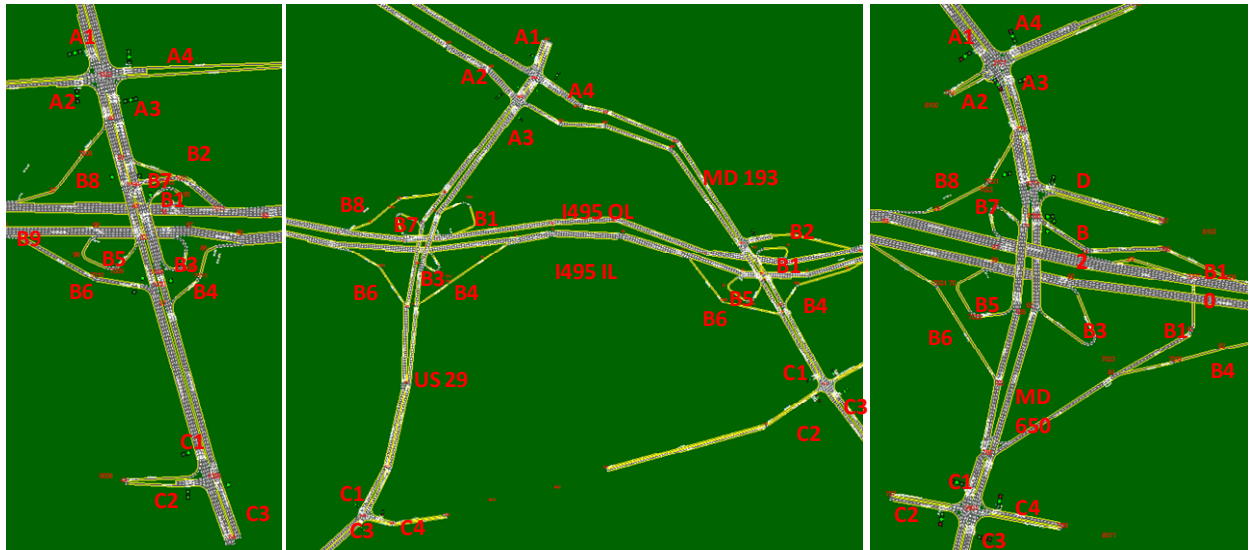
Figure 7-1 Case study site sketch for the multi-interchange control model

7.5.2 Traffic demand pattern

The traffic demands based of year 2011 for each interchange are listed in Table 7-1:

Table 7-1 Demands for the case study network (vehicle per hour)

Approach	Movements	Arterial			
		MD 97	US 29	MD 193	MD 650
A1	Left	2	--	--	99
	Through	2806	3178	--	2320
	Right	105	118	--	19
A2	Left	56	304	--	17
	Through	134	1006	--	10
	Right	259	220	--	60
A3	Left	1	--	--	165
	Through	1099	1509	--	1754
	Right	259	123	--	73
A4	Left	486	354	--	784
	Through	331	1152	--	20
	Right	59	80	--	127
Ramps	B1	840	241	803	301
	B2	337	--	399	636
	B3	321	1007	--	535
	B4	465	527	260	1089
	B5	624	--	511	663
	B6	1008	210	554	441

	B7	468	597	48	632
	B8	770	1399	--	392
C1	Left	--	42	455	63
	Through	3341	2988	530	3198
	Right	135	--	565	123
C2	Left	324	--	34	178
	Through	--	--	269	6
	Right	353	--	79	88
C3	Left	--	--	332	10
	Through	2633	1041	1644	2013
	Right	--	82	14	19
C4	Left	--	307	134	127
	Through	--	--	1895	3
	Right	--	115	10	171
I-495	EB	7703	--	--	--
	WB	--	--	--	7907

To better evaluate the performance of the proposed model, we have developed three different demand scenarios based on the actual traffic demand of the year 2011, namely the Low, Medium, and High demand scenarios. Among them, the Medium demand scenario has actual traffic demand of the year 2011. The Low and High demand scenarios have 90% and 110% of the actual demand, respectively, with the traffic pattern remaining unchanged.

7.5.3 Signal timing optimization and performance comparison methods

The signal plans generated from the proposed model are compared with those generated by Transyt-7F (release 10), which is one of the most advanced programs for both research and practice. Transyt-7F (release 10) offers two optimization algorithms, namely the hill-climbing algorithm and the GA algorithm. The GA method in Transyt-7F (release 10) has been employed to optimize signal timings for the case network, and the proposed model is solved with the hybrid Simulated Annealing Genetic Algorithm (SA-GA) proposed in Chapter 8. Both algorithms run for 50 generations with a population size of 30. The GA in Transyt-7F has a

crossover probability of 0.4, and a mutation probability of 0.01. To take the advantage of the proposed SA-GA algorithm in Chapter 8, the GA for the proposed model has a crossover probability of 0.7, and a mutation probability of 0.5. During the optimization process, the entire network is divided into three signal groups, namely the MD 97 interchange group, the US 29 and MD 193 interchange group, and the MD 650 interchange group. Each signal group has its own cycle length and offsets. All the simulation runs in the signal optimizers are performed for 15 minutes, as recommended in Highway Capacity Manual 2000 (HCM 2000). A network initialization process of 3 minutes is used for all programs.

For comparing the performance fairly, we employ CORSIM to generate MOEs. For the CORSIM simulation model, we calibrate its O-D (origin-destination) tables based on the observed demand pattern. At the link level, free flow speed, physical geometry, and saturated flow for each lane-group have been calibrated. The following statistical results are based on one-hour simulation runs.

7.5.4 Resulting signal timings

Table 7-2 Optimized cycle length of the case study network

Demand Scenario	Model	Signal Group# (second)		
		MD 97	US 29 &MD 193	MD 650
Low	T7F	101	177	177
	MICM	157	100	123
Medium	T7F	148	201	177
	MICM	168	112	147
High	T7F	177	124	210
	MICM	182	112	152

Table 7-2 summarizes the optimized cycle length for all three signal group in the control boundary from both Transyt-7F (T7F) model and the proposed multi-interchange control model (MICM). The cycle length increases with the demand level for both Transyt-7F model and the

99

proposed model for the signal group of MD 97 and MD 650. For the same demand level, the proposed model is intended for shorter cycles than the Transyt-7F model for the MD 650 and US 29 & MD 193 signal group, but longer cycles for the MD 97 group.

7.5.5 Overall system performance comparison

The simulation results from CORSIM for one hour are presented in this section. For each scenario, the average network-wide total delay, total queue delay, and system throughput of 50 simulation runs are listed in Table 7-3. These results in Table 7-3 indicate that the proposed model outperforms Transyt-7F for all three volumes with respect to the average delay and total queue delay. The 95% confidence intervals indicate that the improvements are statistically significant. For the low demand scenario, the proposed model yields 35% less total delay but 2% lower throughputs. For the medium demand scenario, the proposed model produces 20% less total delay and 4% higher throughputs. For the high demand scenario, the proposed model results in 2% more delay and 11% higher throughputs. We can predict that the Transyt-7F timings will yield more delay if counting the delay of those vehicles waiting to enter network, which CORSIM does not count in the total delay.

Table 7-3 Overall model performance comparison

Demand Scenarios		Simulation Results from CORSIM (One hour)				
		MICM	T7F	Improvement*	Improvement*(%)	95% CI*
Low	Total Delay (vehicle-hours)	772.7	1194.9	422.2	35%	[414.0, 430.4]
	Total Queue Delay (vehicle-hours)	418.6	978.9	560.3	57%	[556.0, 564.6]
	Total Throughput (vehicles)	33928	34543	-615	-2%	[-822.6, -407.0]
	Average Delay (second / vehicle)	82	125	42.6	34%	[41.8, 43.4]
Medium	Total Delay (vehicle-hours)	1039.5	1297.7	258.2	20%	[249.4, 267.0]
	Total Queue Delay (vehicle-hours)	590.9	1059.6	468.7	44%	[464.5, 472.9]
	Total Throughput (vehicles)	36841	35349	1492	4%	[1291.2, 1692.6]

	Average Delay (second / vehicle)	102	132	30.6	23%	[29.7, 31.5]
	Total Delay (vehicle-hours)	1377.7	1348.0	-29.7	-2%	[-42.5, -16.9]
High	Total Queue Delay (vehicle-hours)	738.0	1068.2	330.1	31%	[324.2, 336.1]
	Total Throughput (vehicles)	38918	35031	3887	11%	[3727.7, 4045.9]
	Average Delay (second / vehicle)	127	139	11.1	8%	[9.7, 12.5]

* Delay improvement = T7F Delay – MICM Delay

Throughput improvement = MICM Throughput - T7F Throughput

Delay Improvement (%) = (T7F Delay – MICM Delay) / T7F Delay × 100%

Throughput Improvement (%) = (MICM Throughput - T7F Throughput) / T7F Throughput × 100%

95% CI = 95% confidence Interval for the improvement

Average Delay = Total Delay / Total Throughput

Queue delay = Delay calculated by taking vehicles having acceleration rates below 2 feet per second^2 and speeds below 9 feet per second. If a vehicle's speed is below 3 feet per second, it will be included every second. Otherwise it will be included every two seconds (ITT Industries 2006).

7.5.6 Delay comparison by corridor

Tables 7-4 and 7-5 present the total delays and throughputs for the ten major corridors in the network, namely I-495 IL, I-495 OL, MD 650 NB, MD 650 SB, MD 193 EB, MD 193 WB, MD 97 NB, MD 97 SB, US 29 NB, and US 29 SB. Those results indicate that the proposed model yields better performance for the freeways (I-495 IL and I-495 OL). For instance, the proposed model produces almost the same throughputs as the Transyt-7F model (14785 vs. 14882 vehicles) but far less delay (81.3 vs. 105.5 vehicle hours) at the low demand level, less delay (103.3 vs. 108.6 vehicle hours) and more throughputs (16162 vs. 15307 vehicles) at the medium demand level, and more delay (138.7 vs. 108.2 vehicles hours) but far higher throughputs (17344 vs. 15477 vehicles) at the high demand level.

For the arterials, the results are mixed, which implies that the proposed model improves the overall system performance by balancing the different corridors' performance. For instance, at the high demand level, the proposed model yields less delay (102.4 vs. 115.3 vehicle hours) and more throughputs (5143 vs. 4622) than the Transyt-7F model for US29. However, the proposed model produces more delay (132.7 vs. 85.8 vehicle hours) and lower throughputs (5287 vs. 5933 vehicles) for MD 650 at the same demand level.

Table 7-4 Total delay by corridor (vehicle minutes)

Demand Scenarios		Simulation Results from CORSIM (One hour)				
		MICM Model	T7F	Improvement*	Improvement * (%)	Improvement (95% CI)
Low	I-495 IL	34.8	46.7	11.9	26%	[10.8, 13.1]
	I-495 OL	46.5	58.8	12.3	21%	[11.8, 12.8]
	MD 650 NB	27.9	37.3	9.4	25%	[9.0, 9.8]
	MD 650 SB	58.3	45.5	-12.8	-28%	[-14.6, -11.0]
	MD193 EB	30.1	121.6	91.5	75%	[87.7, 95.2]
	MD193 WB	104.8	141.1	36.3	26%	[33.9, 38.7]
	MD97NB	37.9	45.4	7.5	16%	[7.0, 8.0]
	MD97SB	122.2	286.4	164.2	57%	[159.6, 168.8]
	US29 NB	13.6	35.4	21.8	62%	[21.1, 22.6]
	US29 SB	81.9	97.6	15.8	16%	[15.5, 16.0]
Medium	I-495 IL	44.0	48.1	4.1	9%	[3.0, 5.2]
	I-495 OL	59.3	60.5	1.2	2%	[0.6, 1.8]
	MD 650 NB	39.5	37.0	-2.5	-7%	[-3.0, -2.1]
	MD 650 SB	50.1	44.5	-5.6	-13%	[-7.0, -4.2]
	MD193 EB	46.7	148.6	101.9	69%	[98.6, 105.2]
	MD193 WB	104.4	150.9	46.5	31%	[44.6, 48.4]
	MD97NB	42.3	53.3	11.1	21%	[10.7, 11.5]
	MD97SB	254.8	289.6	34.8	12%	[30.1, 39.6]
	US29 NB	15.6	29.0	13.4	46%	[13.0, 13.9]
	US29 SB	83.9	124.0	40.1	32%	[39.7, 40.5]
High	I-495 IL	58.3	49.0	-9.3	-19%	[-11.8, -6.8]
	I-495 OL	80.3	59.2	-21.2	-36%	[-21.9, -20.4]
	MD 650 NB	68.0	26.9	-41.1	-153%	[-42.7, -39.5]
	MD 650 SB	64.7	59.0	-5.7	-10%	[-6.7, -4.7]
	MD193 EB	80.7	99.5	18.9	19%	[14.7, 23.1]
	MD193 WB	106.3	107.0	0.7	1%	[-0.2, 1.5]
	MD97NB	48.5	86.4	37.9	44%	[37.5, 38.3]
	MD97SB	317.6	294.0	-23.5	-8%	[-27.1, -20.0]
	US29 NB	17.9	18.4	0.5	3%	[0.2, 0.9]
	US29 SB	84.5	96.9	12.4	13%	[12.1, 12.7]

* Delay improvement = T7F Delay – MICM Delay; CI = confidence Interval
Delay Improvement (%) = (T7F Delay – MICM Delay) / T7F Delay × 100.

For all three demand levels, the proposed model provides better performance than Transyt-7F in terms of the total system queue delay and average delay. The improvement tends to increase with the congestion level. That is the advantage of optimizing the signal timings with the adjacent interchanges considered in the same model framework. By tackling the traffic dynamics the adjacent interchanges, the proposed model reduces the total delay experienced by the traffic in the network. The results demonstrate that the proposed model is promising.

Table 7-5 Throughput by corridor (vehicle)

Demand Scenarios		Simulation Results from CORSIM (One hour)				
		MICM Model	T7F	Improvement*	Improvement * (%)	Improvement (95% CI)
Low	I-495 IL	6635	6114	521	9%	[519, 523]
	I-495 OL	8150	8768	-618	-7%	[-619, -618]
	MD 650 NB	2118	2355	-237	-10%	[-237, -237]
	MD 650 SB	3170	3577	-408	-11%	[-408, -408]
	MD193 EB	602	510	92	18%	[92, 93]
	MD193 WB	796	873	-77	-9%	[-78, -77]
	MD97NB	1560	1689	-130	-8%	[-130, -129]
	MD97SB	2884	2773	111	4%	[110, 111]
	US29 NB	1675	1527	149	10%	[148, 149]
	US29 SB	2852	2983	-131	-4%	[-131, -130]
Medium	I-495 IL	7297	6448	849	13%	[847, 851]
	I-495 OL	8864	8859	5	0%	[5, 6]
	MD 650 NB	2364	2354	11	0%	[11, 11]
	MD 650 SB	3529	3555	-26	-1%	[-27, -26]
	MD193 EB	652	526	127	24%	[127, 127]
	MD193 WB	827	903	-76	-8%	[-76, -76]
	MD97NB	1711	1701	11	1%	[10, 11]
	MD97SB	3002	2866	136	5%	[136, 137]
	US29 NB	1864	1623	241	15%	[241, 242]
	US29 SB	3002	3040	-38	-1%	[-38, -37]
High	I-495 IL	7807	6632	1174	18%	[1172, 1176]
	I-495 OL	9537	8845	692	8%	[692, 693]

MD 650 NB	2562	2364	198	8%	[197, 198]
MD 650 SB	3623	3224	399	12%	[398, 399]
MD193 EB	654	546	108	20%	[108, 108]
MD193 WB	845	905	-60	-7%	[-61, -60]
MD97NB	1863	1729	134	8%	[134, 135]
MD97SB	3129	2923	206	7%	[206, 206]
US29 NB	1989	1683	307	18%	[306, 307]
US29 SB	3154	2940	214	7%	[214, 214]

* Improvement = MICM Throughput - T7F Throughput

Delay Improvement (%) = (MICM Throughput - T7F Throughput) / T7F Throughput × 100%

CI = confidence Interval

7.6. _Conclusions_

This chapter proposed the capacity reduction model for freeway though traffic with exit queue in the neighboring lane. The proposed model is imbedded into the multi-interchange control model to optimize the signal timings for several adjacent interchanges. Three adjacent interchanges in Silver Spring, MD are employed to demonstrate the performance of the proposed model. The comparison results with Transyt-7F model have demonstrated that signal timings resulting from the proposed model are far better than those with Transyt-7F under a wide range of traffic conditions, especially at high traffic volumes.

Chapter-8:   Hybrid Simulated Annealing and Genetic Algorithm for System-wide Signal Timing Optimization

8.1. *Introduction*

This chapter explores the possibility of developing efficient solution algorithms for the proposed control models. In the previous chapters, all the models are solved with a GA (Genetic Algorithm)-based algorithm, as described in Section 4.3.3. A GA has several limitations:

1. Premature convergence. This is typically the result of the extreme reliance on crossover operations. The dominance of crossover can result in stagnation as the population becomes more homogeneous. If the mutation rate is too low, a GA may experience difficulty in switching to other search areas.

2. Poor local search performance. This leads a GA to a near optimal solution, from which it has difficulty converging to the optimal solution.

3. Large memory use. Since a GA should maintain a large population of solutions, it consumes much memory when the problem dimensions are high.

To improve the performance of a GA for oversaturated signal optimization, we apply the hybrid SA-GA algorithm proposed by Adler (1993) to improve the crossover and mutation operation based on the Metropolis selection rule. The rest of this chapter is organized as follows: the next section presents the problem formulation for oversaturated signal optimization, followed by the decoding scheme for both GA and SA. The details of both the SA and SA-GA hybrid algorithms are introduced in the following two sections. Next, a numerical case study is employed to compare the performance of SA, GA, and SA-GA algorithms. Finally, the last section presents the conclusions.

8.2. *Signal timing encoding scheme*

The oversaturated signal optimization models usually employ the National Electrical Manufacturers Association (NEMA) Eight-Phase signal timing structure to represent the signal of a typical intersection. For a typical four-leg intersection, the NEMA Eight-Phase structure can be illustrated by Figure 4-11, and can be modeled with the following equations:

$$g_{k1} + g_{k2} = g_{k5} + g_{k6} \tag{8-1}$$

$$g_{k3} + g_{k4} = g_{k7} + g_{k8} \tag{8-2}$$

$$g_{k1} + g_{k2} + g_{k3} + g_{k4} = C_k \tag{8-3}$$

$$C_k = C/2^{h_k} \tag{8-4}$$

$$h_k = \begin{cases} 1, signal\ k\ has\ half\ common\ cycle\ length \\ 0, \qquad\qquad otherwise \end{cases} \tag{8-5}$$

$$g_{kj} \geq MG_{kj}, j = 1, \dots, 8 \tag{8-6}$$

$$MinC \leq C_k \leq MaxC \tag{8-7}$$

$$0 \leq offset_k < C_k \tag{8-8}$$

$$g_{kj}, C_k, offset_k\ are\ integers \tag{8-9}$$

in which g_{kj} is the green time for phase j of signal k, C_k is the cycle length of signal k; MGkj is the minimum green time of signal k phase j; MinC is the minimum cycle length; MaxC is the maximum cycle length; C is the common signal cycle length; h_k is a binary variable that indicates whether signal k has a half common cycle length as defined by Equation (8-5); and offset$_k$ represents the offset of signal k. Equations (8-1) and (8-2) indicate the existence of the signal barrier. Equations (8-3) and (8-4) enforce the cycle length constraints. Equation (8-6) restricts the green time of each phase to at least its minimum green time, and Equation (4-28)

specifies user-defined minimum and maximum cycle lengths. Equation (8-8) requires that the offset of signal k lie between 0 and its cycle length.

The objective function of a signal optimization model for an oversaturated intersection may maximize the system throughput or minimize total system delay. The system throughput or system delay is computed by some macroscopic, mesoscopic, or microscopic traffic flow models, which can capture link or lane blockage, or both of them. This study employs the Enhanced Cell Transmission Model (CTM) proposed in Chapter 4 to compute the traffic Measure of Effectiveness (MOE). The objective function minimized in this model represents total system delay:

$$Min \quad Z = \sum_{t=0}^{T} \sum_{i} w_i (n_i^t - \sum_{j \in \Gamma(i)} y_{ij}^t \, y_{ij}^t) \qquad (8\text{-}10)$$

where w_i is a weighted coefficient to modify the relative importance of each cell, $\Gamma(i)$ is the downstream cell set of cell i, and T is total operation time period, and n_i^t denotes the number of vehicles in each cell at time t. More details can be found in Chapter 4.

8.3. *SA algorithm*

Annealing is the cooling process through which a low energy state is reached in a solid. Its main steps can be described as follows: First, the solid is heated to melt at a high temperature, in which state all particles arrange randomly. Then, the temperature is slowly lowered until the particles arrange themselves in their minimum energy state, which is the ground state. To prevent the resulting crystals from having defects or even lacking all crystalline order, the cooling schedule should be sufficiently slow.
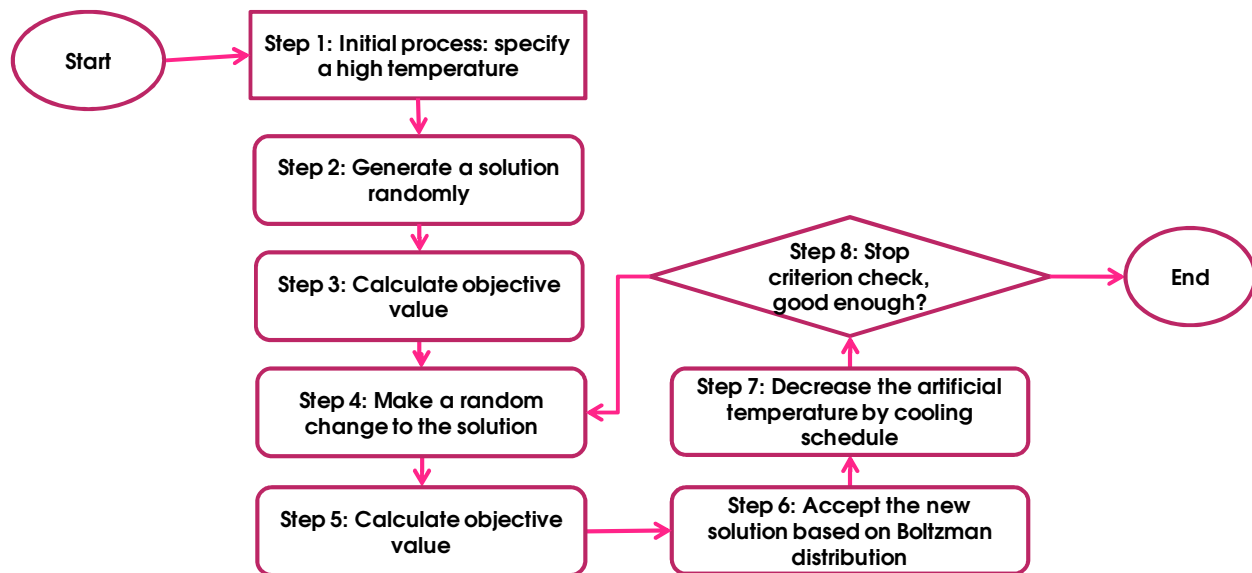
Figure 8-1 Simulated Annealing (SA) algorithm flow chart

The analogy between the optimization problem and a solid is based on the following two aspects: the solutions of the optimization problem are equivalent to the states of a solid, and the objective value is equivalent to the energy of a state. For the oversaturated signal timings optimization problem, the solutions are signal timings and the objective value could be any MOE (Measure of Effectiveness) of the signal timings.

The flow chart of the SA algorithm can be summarized in Figure 8-1. The detailed algorithmic steps are as follows:

Step 1. Specify a high initial artificial temperature T_0;

Step 2. Randomly generate an initial solution S_1 and let n=1;

Step 3. Compute the objective function of the initial solution Z_1;

Step 4. Make a random change to solution S_n and obtain a new solution S_{n+1};

Step 5. Compute the objective function value Z_{n+1} of the new solution S_{n+1};

Step 6. If $Z_{n+1} > Z_n$, the probability of accepting these solutions is computed from the

  Boltzmann distribution as follows:

108

$$P(S_{n+1}) = \frac{1}{b} \exp\left(-\frac{Z_{n+1} - Z_n}{T_{n+1}}\right) \qquad (8\text{-}11)$$

where $P(S_{n+1})$ is the probability of accepting the new solution; T_{n+1} is the

current artificial temperature, and b is a normalization constant. This acceptance

rule is referred as the Metropolis criterion (Kirkpatrick 1984). The Metropolis

selection procedure is as follows:

a. Generate a random number r according to uniform distribution between 0 and 1;

b. If $P(S_{n+1}) > r$, accept the new solution; otherwise reject the new solution,

$S_{n+1} = S_n$ and $Z_{n+1} = Z_n$.

Step 7. Decrease the artificial temperature according to the cooling schedule, $T_{n+1} =$

$f(T_n, T_0)$;

Step 8. Check if the stop criterion is satisfied. If yes, stop and return S_{n+1}. Otherwise, let

n=n+1 and go back to Step 4.

In this study, the random change in Step 4 is done by the GA's mutation operation.

8.3.1 Cooling schedule

In the annealing schedule, an initial high artificial temperature is provided and then the

temperature is slowly lowered through successive iterations. The process should be slow enough

to allow sufficient time for the state to reach equilibrium at each temperature. This study

employs the following well-known cooling schedule that provides necessary and sufficient

conditions for convergence (Hwang and He 2006):

$$T(t) = \frac{T_0}{\ln(t)} \quad \forall t > 0 \qquad (8\text{-}12)$$

where T(t) is the artificial temperature at time t, T_0 is the initial temperature, and t denotes the

time. When t approaches infinity, T(t) approaches zero. In this study, time t is represented by the discrete time step n. Equation (8-12) can be transformed to the following:

$$T(n) = \frac{T_0}{\ln(n)} \quad \forall \, n > 0 \tag{8-13}$$

The acceptance probability can be converted to Equation (8-14)

$$P(S_{n+1}) = \frac{1}{b} \exp\left[-\frac{Z_{n+1} - Z_n}{T_0} \log(n+1)\right] = \frac{1}{b}(n+1)^{-\frac{Z_{n+1}-Z_n}{T_0}} \tag{8-14}$$

Then

$$P(S_n) = \frac{1}{b} n^{-\frac{Z_n - Z_{n-1}}{T_0}} \tag{8-15}$$

Since $P(S_0) = 1$, the normalization parameter b can be set to 1 to satisfy for this cooling schedule. The final acceptance probability can be stated as:

$$P(S_n) = n^{-\frac{Z_n - Z_{n-1}}{T_0}}, \forall n \geq 1 \tag{8-16}$$

When a new solution from the neighborhood of the current solution is generated, if its objective value is better (less) than that of the current solution, it is accepted. If its objective value is worse (larger) than that of the current solution, it is accepted based on Equation (8-16). The probability determined by Equation (8-16) depends on the initial artificial temperature T_0. However, it is challenging to determine T_0 since it depends on the strategies for solving the problem. In general, T_0 is a function of the maximum and minimum objective function value. In this study, T_0 is set to the absolute value of the objective function value of the initial solution, i.e., $T_0 = |Z_0|$.

8.4. *A SA-GA hybrid algorithm*

The hybrid SA-GA algorithm proposed by Adler (1993) is adapted to improve the

crossover and mutation based on Metropolis selection rule. To combine SA with GA, this method replaces the mutation and recombination operators with SA-Mutation (SAM) and SA-Recombination (SAR). Similarly to a standard mutation operator, the SAM operator mutates a solution and returns it. However, SAM evaluates the fitness of the solution after mutation, and decides whether to accept it or just stay with the previous one, based on the Metropolis rule. The SAR operator is also similar to the standard recombination operator. It generates two child solutions based on two parent solutions. During the SAR, the crossover is applied to generate two children. Each of the children is then compared to the better parent for acceptance by the aforementioned Metropolis rule. The artificial temperature will decrease in each generation according the cooling schedule. The final SA-GA algorithm is summarized in Figure 8-2.
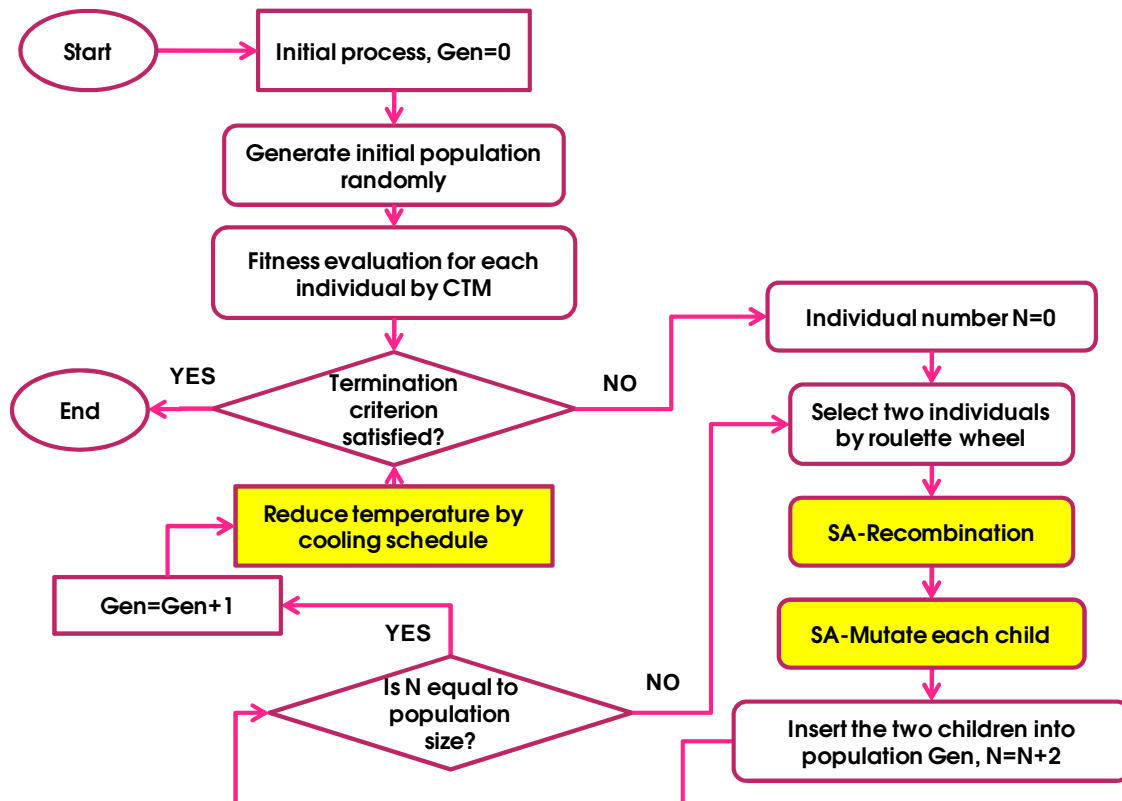
Figure 8-2 SA-GA hybrid algorithm flow chart

The detailed steps of SA-GA algorithm are as follows:

Step 1. Initialize by setting generation iterator Gen=0;

Step 2. Generate the initial population randomly;

Step 3. Evaluate the fitness of each individual in the initial population by CTM;

Step 4. Check if the termination criterion is satisfied. If yes, terminate and return the best

solution; otherwise, reset the individual number of current generation to zero (set

N=0) and continue to the next step;

Step 5. Select two individuals by roulette wheel from the current population;

Step 6. SA-Recombine the parents to generate two child solutions;

Step 7. SA-Mutate each child;

Step 8. Insert the two children into the new population, and set N=N+2;

Step 9. Check if the new population has enough individuals. If yes, set Gen=Gen+1,

reduce the artificial temperature according to the cooling schedule and go to Step

4; otherwise go to Step 4.

8.5. *Numerical case study*

To evaluate the performance of the proposed algorithms, a segment of Georgia Avenue (MD97) intersecting with the Capital Beltway in Silver Spring, Maryland has been selected for the experimental study. As shown in Figure 6-1, the target site includes four signalized intersections from Forest Glen Rd (MD192) to Seminary Place. The actual volume of each approach is used for performance evaluation (see Table 6-1).

To compare the performance of the three algorithms, SA, GA, and SA-GA are implemented in C# and applied to solve the model presented above using the same data structure on a desktop with an six-core Intel(R) Xeon(R) CPU (X5650 @ 2.67GHz) and 6 GB RAM. Both GA and SA-GA optimizers run for 200 generations with a population size of 30 (SA runs for

3000 iterations), a crossover probability of 0.3, a mutation probability of 0.2, and the same random seed 3. 15 minutes of traffic movements with signal optimization are simulated in all the simulation runs, as recommended in the Highway Capacity Manual 2000 (HCM 2000).

Table 8-1 Demands for the case study site

Entrance	Movements	Demand (vehicle per hour)
A	Through	3,382
	Right	112
B	Left	44
	Through	101
	Right	179
C	Left	596
	Through	340
	Right	47
D	Left	315
	Right	227
E	Right	1,200
F	Right	350
G	Left	553
	Right	25
H	Through	2,715
Total	-----	10,186

Figure 8-3 illustrates the evolution of the objective function values of SA iterations over CPU time. SA uses 149.81 seconds of CPU time to complete all 3000 iterations. Figure 8-3 indicates how SA gradually reduces the objective function value, although the current solutions objective value frequently jumps to a high value. After 117.17 seconds of CPU time, SA has reached its best solution with an objective value 42.20.

Figure 8-4 and Figure 8-5 represent the evolution of the objective function value over CPU time for the GA and SA-GA hybrid algorithms. They show the minimum, average, and maximum of the objective function values in each generation. For the 200th generation, CPU times are 127.25 seconds for GA and 147.81 seconds for SA-GA. The SA-GA hybrid algorithm requires 20.56 seconds (14%) more than GA. To run the same number of generations, SA-GA needs more CPU time than GA because SA-GA requires additional time to evaluate each

chromosome during crossover. It can be expected that as the crossover rate increases, SA-GA will use more CPU time.
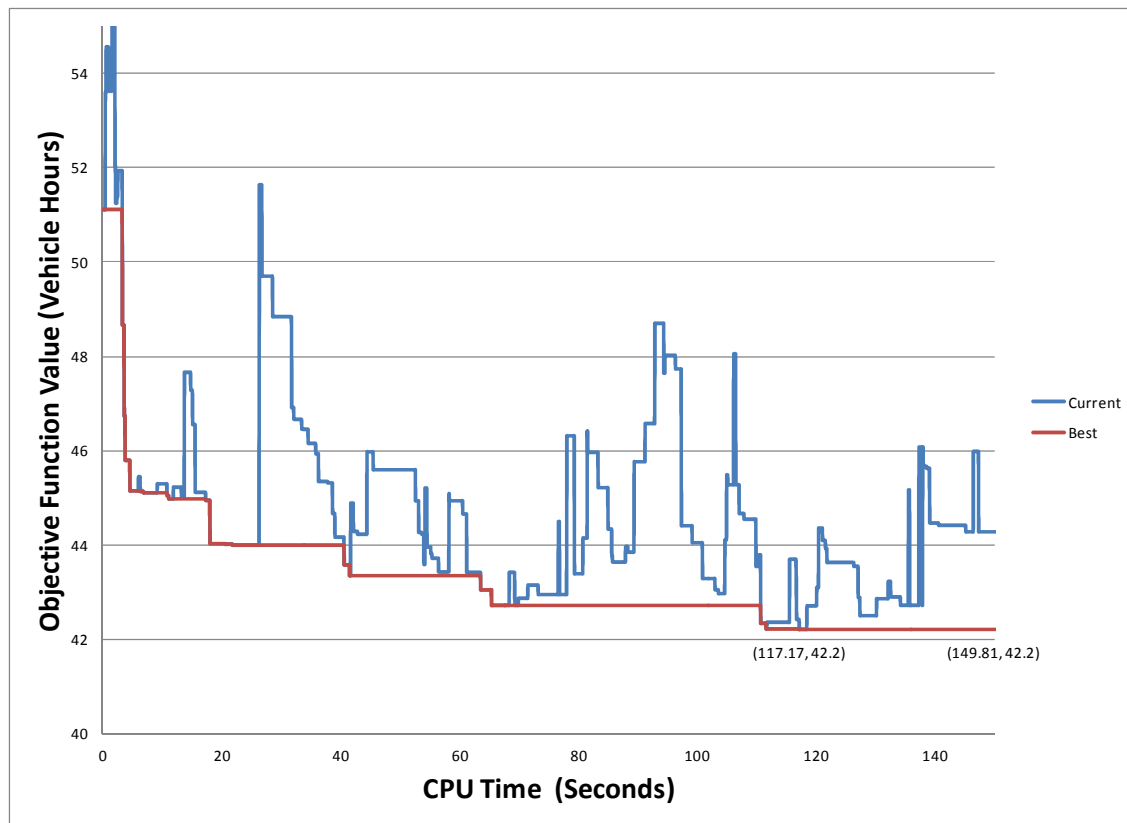

Figure 8-3 Evolution of objective function value over CPU time for SA

Both GA and SA-GA employ the elitist selection method. That can explain why the minimum objective values of each generation decrease steadily. Both the average and maximum values of the objective function exhibit fluctuations for those two algorithms. However, the generation average for SA-GA is smoother than for GA, due to SA-Mutation and SA-Recombination operations. It is expected that the generation average for SA-GA will be smoother at later generations. This is due to the Metropolis rule since the acceptance probability determined by Equation (8-16) decreases with additional generations, which suggests that these solutions with larger objective values have lower probabilities of being accepted.
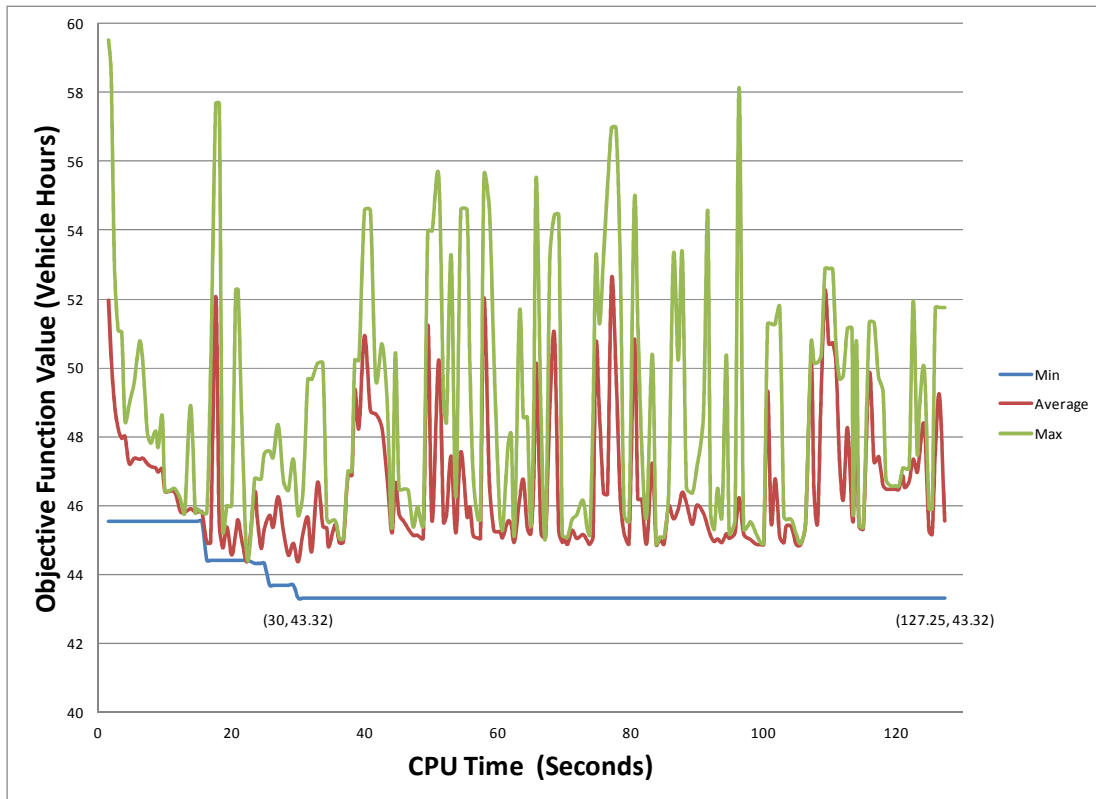
114

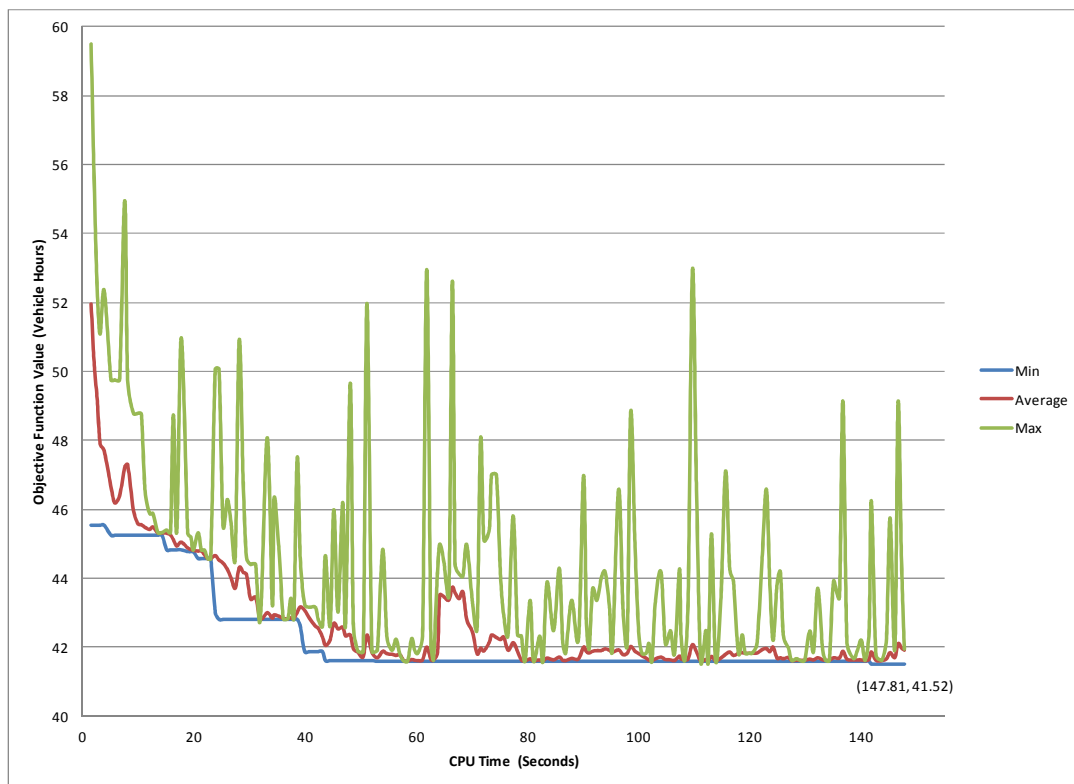Figure 8-4 Evolution of objective function value over CPU time for GA

Figure 8-5 Evolution of objective values over CPU time for SA-GA hybrid algorithm

115

Figure 8-6 compares the performance of SA, GA, and SA-GA. It indicates that SA-GA outperforms both GA and SA in terms of finding a better solution or finding a comparative good solution in a shorter time. GA finds its best solution after 30.00 seconds of CPU time with objective value 43.32. SA-GA reaches a comparative good solution after 27.44 seconds of CPU time (35 iterations) with objective value 42.71. SA finds a comparative good solution with objective value 43.3 after 41.5 seconds, and its optimized solution after 117.17 seconds with objective value 42.20. SA-GA obtains the first comparative good solution as SA after 39.94 seconds with objective value 41.88, and its optimized solution after 141.87 seconds with objective value 41.52. It is notable that the difference between the objective values of the optimized solutions of the three algorithms, which is 1.78 between those of GA and SA-GA, and 0.68 between SA and SA-GA. However, SA is still preferable in this case study since it can find a better solution and reach a comparative good solution with less computation time.
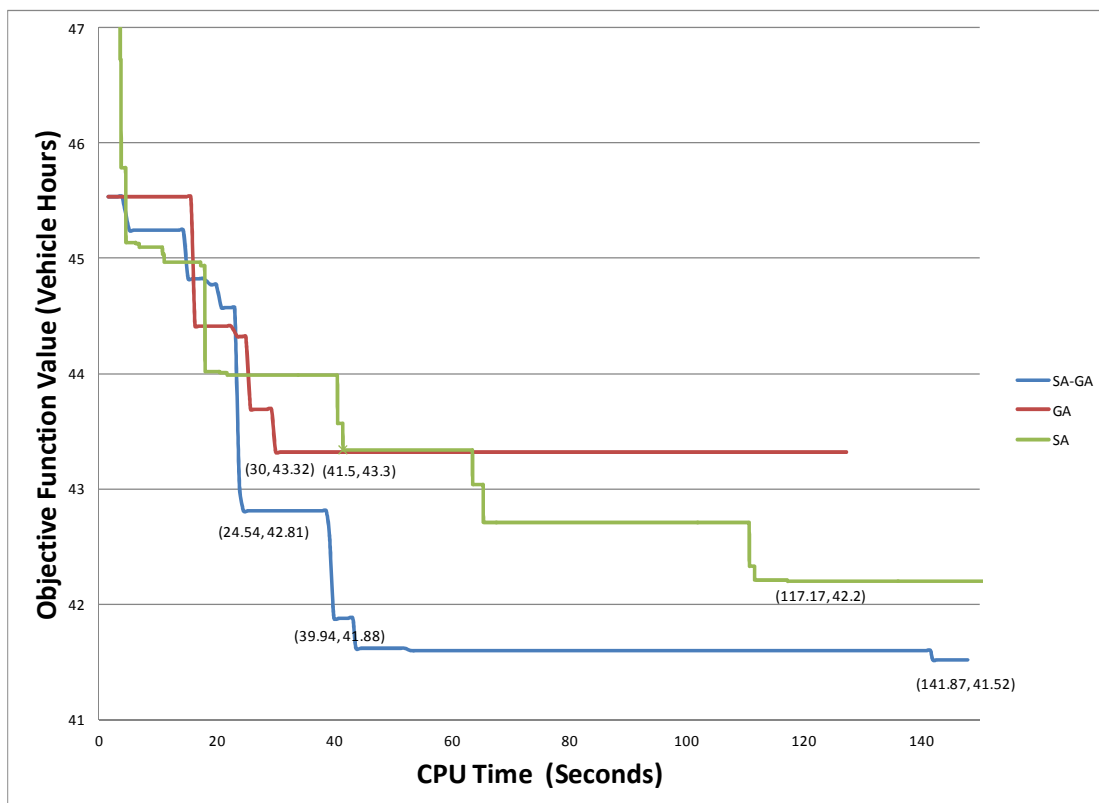

Figure 8-6 Performance comparison of SA, GA, and SA-GA

116

All the three algorithms are known to be unstable, which means their optimized solutions may fluctuate with different random seeds. To test their stability, we run them each for 50 times with different random seeds and list the estimated sample mean, sample standard deviation (STD), 95% confidence interval (95% CI) of sample, and 95% confidence interval (95% CI) of sample mean in Table 8-2.

Table 8-2 Stability test for SA, GA, and SA-GA

Algorithm	Mean	STD	95% CI of sample	95% CI of sample mean
SA	44.81	1.98	[40.92, 48.69]	--/--
GA	44.71	1.03	[42.69, 46.73]	--/--
SA-GA	44.04	1.18	[41.73, 46.35]	--/--
GA – [SA-GA]	0.67	1.01	--/--	[0.39, 0.95]
SA – [SA-GA]	0.76	0.80	--/--	[0.54, 0.99]

The statistical results presented in Table 8-2 indicate that the optimized solutions of SA-GA have the lowest mean, and those of SA have the highest mean. Those differences are significant, as indicated by the 95% CI of sample means of GA – [SA-GA] and SA- [SA-GA]. The standard deviation measures the algorithm stability. GA has the lowest standard deviation and SA-GA has a close one. For SA-GA, on 95% confidence level, we will find a solution between 41.73 and 46.35. The difference between the best solution and the worst is 5.62, which equals to 5.62 vehicle hour delay for a 15 minutes period. It is accepted for traffic operation purposes.

It would be desirable to check the goodness of the optimized solution by comparing it with the exact optimal solution. However, we do not know the exact optimal solution for those problems since they are nonlinear integer programming problems. Normally, we could obtain a lower bound for such nonlinear integer programming problems by relaxing the nonlinear constraints to linear constraints, eliminating the integer requirements, and working with a small

117

problem. However, these strategies are still very difficult to apply in our models due to their problem size. In the proposed models, the smallest problem deals with arterials having several signalized intersections since we deal with oversaturated traffic conditions and must handle spillback between adjacent intersections. In CTM, for each time interval, we have 8 integer variables for the timing of each signal, and 3 variables for each ordinary cell and many more variables for merging and diverging cells. To model the traffic flow with sufficient details, the time period should not exceed 10 seconds for arterial modeling. For the network using for this numerical case study, there are about 40 cells if the time period is set to 10 seconds. Therefore, for fifteen minutes, we have 90 time periods and more than 14400 variables in the model. Among them, 720 are integer variables. In addition to the variable number, some of the constraints of the proposed models' are nonlinear, which are very difficult to linearize, e.g. Equation (4-16) and (5-17). All those difficulties make lower bounds of the optimization problems very hard to obtain. Therefore, we test the goodness of the optimized solution with an indirect method described below.

In the rest of this section, we conduct a sampling experiment to check the algorithm's goodness indirectly using the method used for a GA by Jong and Schonfeld (2003). In this experiment, we first randomly generate 10,000 samples and compute their objective values, thus obtaining an approximation of the actual distribution of system total delays for all possible signal timings in the feasible domain. The 50,000 samples range from 41.86 to 63.87, have a mean of 51.58 and standard deviation of 3.06. Figure 8-7 illustrates the fitted normal distribution. The optimized solution of SA-GA is 41.25, which is better than the best of the 50,000 samples, with a cumulative probability of $\Pr(F \leq 41.25) = \Pr\left(z \leq \frac{41.25 - 51.58}{3.06}\right) \cong 0.0004$. It indicates that the optimized solution is better than 99.96 percent of solutions in the feasible domain. For SA-GA's

upper limit of 95% sample CI, 46.35, it has a cumulative probability of $\Pr(F \leq 46.35) =$
$\Pr\left(z \leq \frac{46.35-51.58}{3.06}\right) \cong 0.03$, which is still better than 97 percent of solutions.
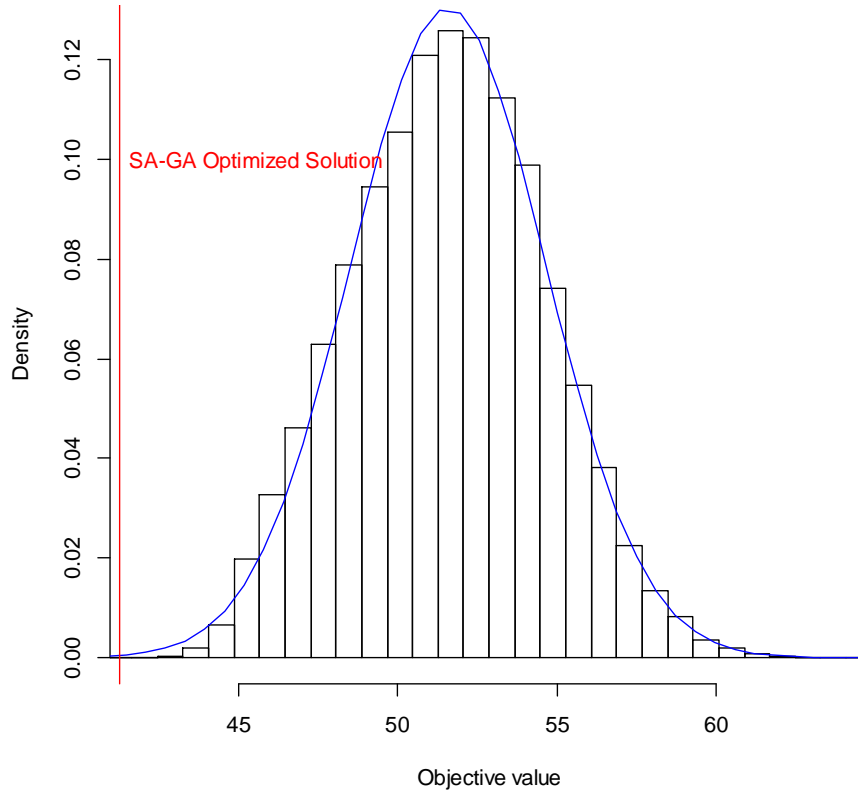
Figure 8-7 The fitted normal distribution of the objective values

8.6. *Conclusions and limitations*

An efficient solution algorithm is a crucial component for implementing a system-wide signal control model, especially for an on-line system. This chapter presents an SA algorithm and an SA-GA hybrid algorithm for solving a system-wide signal optimization model, especially for oversaturated intersections. These algorithms can be easily adapted to a wide-range of signal timing optimization models. To demonstrate the performance of these algorithms, they are employed to solve an optimization model for arterial signal timing under oversaturated

conditions based on an enhanced CTM. The results indicate that the SA-GA hybrid algorithm provides a better optimized solution or a compatible solution in a shorter time. The performance comparison indicates that the SA-GA presented here is promising for solving signal timing optimization problems for interrelated and oversaturated intersections. It should be noted that the algorithms proposed here are stochastic in nature and their performance could change with parameter changes. Further research should explore the combinations of parameters and cooling schedules that are most desirable for various problem characteristics.

Chapter-9:   Conclusions and Future Research Directions

9.1. <u>Research Summary and Contributions</u>

This dissertation focuses on developing an integrated freeway and arterial control model, especially for oversaturated condition. Chapter 1 of this dissertation starts from analyzing the operation issues around a congested interchange. These reveal the need for developing an integrated control model for freeways and arterials to achieve a system-wide optimum rather than their individual performance. Chapter 1 then presents the critical theoretical and operational issues should be addressed in developing such a model, and the objective of this dissertation.

Chapter 2 summarizes a comprehensive review of the relevant studies on both theoretical and practical aspects of an integrated control model of freeways and arterials. After reviewing the limited models directly relative to the integrated control model of freeways and arterials, this review further summarizes the control models that have been developed for freeways or arterial systems. This literature review not only identifies the gap between the state-of-art models and the practical operation needs, but also reveals the promising research direction.

Based on the literature review findings and the operational need of practice, Chapter 3 propose an integrated control model for recurrent congestion. After a detailed discussion of the major research issues and challenges, Chapter 3 presents the control flowchart of the proposed model. Following the logic of the flowchart, Chapter 3 divides the whole model into several principal components with a details description of their functions and interactions.

Chapter 4 deals with the simplest congestion level in this study, in which the congestion is limited to arterials. To capture lane-blockage between adjacent movements, Chapter 4 presents

an enhanced Cell-Transmission Model with the innovative sub-cell concept. Based on the enhanced CTM, Chapter 4 proposes a signal optimization model for arterial signal timings under oversaturated conditions. The proposed signal optimization model is expected to optimize the cycle length, splits, and offsets of adjacent signals along an arterial, under the presence of link-blockage and lane-blockage.

If arterial congestion develops further and spreads itself to freeways, it is essential to consider the freeway delay when optimizing the signals of an interchange and its adjacent intersections. Chapter 5 presents an integrated control model for this congestion level. The proposed model formulates the merging, propagation, and diverging zone of a basic freeway segment with the Cell-Transmission concept. Combing those formulations with the model for arterials presented in Chapter 4, the proposed integrated control model optimizes the traffic signal timings around an interchange with respect to the overall system performance instead of arterials or freeways separately. The proposed model considers the on-ramp and off-ramp spillback, link spillback, and lane blockage simultaneously in an integrated manner.

To demonstrate the performance of the proposed arterial signal optimization model and integrated model for a single interchange, we have conducted extensive numerical experiments in Chapter 6 with the Capital Beltway / Georgia Avenue (MD97) interchange in Silver Spring, MD. In this chapter, we first optimize the signal timings with the proposed models in the network, and then compare their performance with those from Transyt-7F model. To avoid comparison bias, we employ CORSIM model as the performance index provider. The statistical results demonstrate that the proposed models outperform Transyt-7F under a wide range of traffic conditions, especially for high off-ramp volumes. The first half of Chapter 6 presents the results for the proposed arterial signal optimization model. The results indicate that the model

successfully improves system performance by preventing link-blockage and lane-blockage, especially under oversaturated conditions. The second half of Chapter 6 reports the numerical results for the proposed integrated signal interchange model. These results confirm that it is highly desirable to jointly control the freeways and arterials around a congested interchange in order to improve system performance.

In Chapter 7, a multi-interchange control model is proposed to optimize the signals of several arterial corridors connected by a freeway corridor under oversaturated traffic conditions. To capture the effects of exit queue to through traffic on freeway caused by lateral friction, Chapter 7 first proposes a car-following model and derives a capacity deduction model to represent the exit queue effects. The proposed model is then imbedded into the multi-interchange control model to optimize the signal timings for several adjacent interchanges. Four adjacent interchanges in Silver Spring, MD are studied in the numerical experiment to demonstrate its performance. By comparing with Transyt-7F model, we have demonstrated that those signal timings resulting from the proposed model are far better under a wide range of traffic conditions, especially at high traffic volumes.

Chapter 8 presents an SA algorithm and an SA-GA hybrid algorithm for solving a system-wide signal optimization model, especially for oversaturated intersections. These algorithms can be easily adapted to a wide-range of signal timing optimization models. To demonstrate the performance of these algorithms, they are employed to solve an oversaturated signal timings optimization model based on an enhanced CTM. The results indicate that the SA-GA hybrid algorithm provides a better convergence rate and a better optimized solution. The performance comparison indicates that the SA-GA presented here is promising for solving signal timing optimization problems for interrelated and oversaturated intersections. It should be noted

123

that the algorithms proposed here are stochastic in nature and their performance could change with parameter changes. Further research should explore the combinations of parameters and cooling schedules that are most desirable for various problem characteristics.

In summary, the key contributions of this dissertation include:

1. Formulated the lane-blockage problem between the movements of an arterial intersection approach as an linear program with the proposed sub-cell concept, and proposed an arterial signal optimization model under oversaturated traffic conditions;

2. Formulated the traffic dynamics of a freeway segment with cell-transmission concept, and considering the exit queue effects on its neighboring through lane traffic with the proposed capacity model, which is able to take the lateral friction into account;

3. Developed an integrated control model for multiple freeway interchanges, which can capture the off-ramp spillback, freeway mainline spillback, and arterial lane and link blockage simultaneously;

4. Explored the effectiveness of different solution algorithms (GA, SA, and SA-GA) for the proposed integrated control models, and conducted goodness check for the proposed algorithms, which has demonstrated the advantage of the proposed model;

5. Conducted intensive numerical experiments for the proposed control models, and compared the performance of the optimized signal timings from the proposed models with those from Transyt-7F by CORSIM simulations. These comparisons have demonstrated the advantages of the proposed models, especially under oversaturated traffic conditions.

9.2. <u>Future Research Directions</u>

In this study, the calibration and validation of the proposed traffic flow models based on the Cell-Transmission concept remains unfinished due to resource limitations. The procedure for calibrating and validating the proposed models can be performed on the same road segment. The following procedure is based on the method proposed by Muñoz et al. (2004). This method requires traffic data for each cell during each time period under both under-saturated and oversaturated traffic conditions. These data include average speed, average density, inflow from each upstream cell, and outflow to each downstream cell. They can be measured directly or indirectly by detectors or video tapes, or they can be collected manually. With those data, we can perform a least-squares fit on the flow-density curve for each cell. From the resulting flow-density curve, we can obtain the free flow speed, capacity, and the backward propagating speed. We calibrate the proposed model with half of the collected data and validate with the other half. To calibrate the proposed capacity reduction model, we must also record the exit queue length for each time period. By performing a least-squares fit to the capacity reduction model with the collected data, we can estimate its parameters. After we obtain those parameters, we can validate the model by comparing the output flow from the calibrated model with field data.

In addition to the calibration and validation, future studies in this area might be conducted as follows:

Conduct field before-and-after study. The proposed models are available for field implementation. Although we have demonstrated our models' advantages by intensive simulation, it is desirable to conduct before-and-after field studies to verify their benefits.

Develop a detour control model for non-recurrent traffic congestion. A large percentage of traffic delay is caused by non-recurrent traffic congestion. The proposed model has a capacity reduction model for through traffic with a queue in the neighboring lane, which could be adapted to model the capacity reduction with lane closure caused by incidents. With a new component to consider the detour traffic, we can generate emergency traffic control plans to detour traffic to parallel arterials.

References

Abu-Lebdeh, G., and Benekohal, R. (1997). "Development of Traffic Control and Queue Management Procedures for Oversaturated Arterials." *Transportation Research Record: Journal of the Transportation Research Board*, 1603(-1), 119-127.

Abu-Lebdeh, G., and Benekohal, R. (2000). "Genetic Algorithms for Traffic Signal Control and Queue Management of Oversaturated Two-Way Arterials." *Transportation Research Record: Journal of the Transportation Research Board*, 1727(-1), 61-67.

Abu-Lebdeh, G., and Benekohal, R. (2003). "Design and Evaluation of Dynamic Traffic Management Strategies for Congested Conditions." *Transportation Research Part A*, 37(2), 109-127.

Abu-Lebdeh, G., Chen, H., and Benekohal, R. (2007). "Modeling Traffic Output for Design of Dynamic Multi-Cycle Control in Congested Conditions." *Journal of Intelligent Transportation Systems*, 11(1), 25-40.

Adler, D. "Genetic Algorithms and Simulated Annealing: A Marriage Proposal." *Neural Networks, 1993., IEEE International Conference on*, 1104-1109 vol.2.

Ahn, S., Bertini, R., Auffray, B., Ross, J., and Eshel, O. (2007). "Evaluating Benefits of Systemwide Adaptive Ramp-Metering Strategy in Portland, Oregon." *Transportation Research Record: Journal of the Transportation Research Board*, 2012(-1), 47-56.

Allsop, R. (1971). "Sigset: A Computer Program for Calculating Traffic Capacity of Signal-Controlled Road Junctions." *Traffic Eng. Control*, 12, 58–60.

Allsop, R. (1976). "Sigcap: A Computer Program for Assessing the Traffic Capacity of Signal-Controlled Road Junctions." *Traffic Engineering & Control*, 17, 338-341.

Arnold, E. (1998). "Ramp Metering: A Review of the Literature." *Virginia Transportation Research Council Technical Report VTRC99-TAR5*.

Banks, J. (1993). "Effect of Response Limitations on Traffic-Responsive Ramp Metering." *Transportation Research Record*, 1394, 17-25.

Binning, J. C., Burtenshaw, G., and Crabtree, M. (2008). "Transyt 13 User Guide." Transport Research Laboratory, UK.

Bogenberger, K., and May, A. (1999). "Advanced Coordinated Traffic Responsive Ramp Metering Strategies." *California PATH Paper UCB-ITS-PWP-99-19*.

Boillot, F., Blosseville, J. M., Lesort, J. B., Motyka, V., Papageorgiou, M., and Sellam, S. "Optimal Signal Control of Urban Traffic Networks." *Road Traffic Monitoring, 1992 (IEE Conf. Pub. 355)*, 75.

Case, H. W., Hulbert, S. F., Mount, G. E., and Brenner, R. "Effect of a Roadside Structure on a Lateral Placement of Motor Vehicles." *Highway Research Board Annual Meeting*, 364-370.

Cassidy, M. J., Anani, S. B., and Haigwood, J. M. (2002). "Study of Freeway Traffic near an Off-Ramp." *Transportation Research Part A: Policy and Practice*, 36(6), 563-572.

Ceylan, H. (2006). "Developing Combined Genetic Algorithm—Hill-Climbing Optimization Method for Area Traffic Control." *Journal of Transportation Engineering*, 132, 663.

Ceylan, H., and Bell, M. (2004a). "Traffic Signal Timing Optimisation Based on Genetic Algorithm Approach, Including Drivers' Routing." *Transportation Research Part B*, 38(4), 329-342.

Ceylan, H., and Bell, M. G. H. (2004b). "Traffic Signal Timing Optimization Based on Genetic Algorithm Approach, Including Drivers' Routing." *Transportation Research Part B*, 38(4), 329-342.

Ceylan, H., Haldenbilen, S., and Baskan, O. (2010). "Development of Delay Models with Quasi-Newton Method Resulting from Transyt Traffic Model." *Journal of Scientific & Industrial Research*, 69(2), 87-93.

Chang, E., Cohen, S., and Liu, C. (1988). "Maxband-86: Program for Optimizing Left Turn Phase Sequence in Multiarterial Closed Networks." *Transportation Research Record*, 1181, 61-67.

Chang, G.-L., Jifeng, W., and Cohen, S. L. (1994). "Integrated Real-Time Ramp Metering Modeling for Non-Recurrent Congestion: Framework and Preliminary Results." *Transportation Research Record*, 1446, 56-65.

Chang, G., Ho, P., and Wei, C. (1993). "A Dynamic System-Optimum Control Model for Commuting Traffic Corridors." *Transportation Research Part C: Emerging Technologies*, 1(1), 3-22.

Chang, T.-H., and Lin, J.-T. (2000). "Optimal Signal Timing for an Oversaturated Intersection." *Transportation Research Part B*, 34(6), 471-491.

Chang, T., and Sun, G. (2004). "Modeling and Optimization of an Oversaturated Signalized Network." *Transportation Research Part B*, 38(8), 687-707.

Chaudhary, N., Pinnoi, A., and Messer, C. (1991). "Proposed Enhancements to Maxband 86 Program." *Transportation Research Record*, 1324, 98-104.

Chen, C., Cruz, J. B., and Paquet, J. G. (1974a). "Entrance Ramp Control for Travel Rate Maximization in Expressways." *Transportation Research*, 8, 503-508.

Chen, C. I., Cruz, J. B., and Paquet, J. G. (1974b). "Entrance Ramp Control for Travel Rate Maximization in Expressways." *Transportation Research*, 8(6), 503-508.

Chen, L., May, A., and Auslander, D. (1990). "Freeway Ramp Control Using Fuzzy Set Theory for Inexact Reasoning." *Transportation research. Part A: general*, 24(1), 15-25.

Chen, O., Hotz, A., and Ben-Akiva, M. "Development and Evaluation of a Dynamic Ramp Metering Control Model." *8th IFAC(International Federation of Automatic Control)/IFIP/IFORS Symposium on Transportation Systems* Chania, Greece, 1162-1168.

Chlewicki, G. "New Interchange and Intersection Designs: The Synchronized Split-Phasing Intersection and the Diverging Diamond Interchange." *2nd Urban Street Symposium: Uptown, Downtown, or Small Town: Designing Urban Streets That Work*, Anaheim, CA.

Choi, B. (1997). "Adaptive Signal Control for Oversaturated Arterials," Polytechnic University.

Cremer, M., and Schoof, S. (1989). "On Control Strategies for Urban Traffic Corridors." *Proceedings of IFAC/IFIP/IFORS Symposium*, 213-219.

Daganzo, C. F. (1994). "The Cell Transmission Model: A Dynamic Representation of Highway Traffic Consistent with the Hydrodynamic Theory." *Transportation research. Part B: methodological*, 28(4), 269-287.

Daganzo, C. F. (1995). "The Cell Transmission Model. Ii: Network Traffic." *Transportation research. Part B: methodological*, 29(2), 79-93.

Dorothy, P., Maleck, T., and Miller, K. (1998). "Operational Aspects of the Michigan Urban Diamond Interchange." *Transportation Research Record: Journal of the Transportation Research Board*, 1612(-1), 55-66.

Engelbrecht, R., and Barnes, K. (2003). "Advanced Traffic Signal Control for Diamond Interchanges." *Transportation Research Record: Journal of the Transportation Research Board*, 1856(-1), 231-238.

Fang, F., and Elefteriadou, L. (2006). "Development of an Optimization Methodology for Adaptive Traffic Signal Control at Diamond Interchanges." *Journal of Transportation Engineering*, 132, 629.

Farzaneh, M., and Rakha, H. (2006). "Procedures for Calibrating Transyt Platoon Dispersion Model." *Journal of Transportation Engineering*, 132(7), 548-554.

Gartner, N. (1983). "Opac: A Demand-Responsive Strategy for Traffic Signal Control." *Transportation Research Record*, 906, 75-81.

Gartner, N., Assman, S., Lasaga, F., and Hou, D. (1991). "A Multi-Band Approach to Arterial Traffic Signal Optimization." *Transportation Research Part B: Methodological*, 25(1), 55-74.

Gazis, D., Herman, R., and Potts, R. (1959). "Car-Following Theory of Steady-State Traffic Flow." *OPERATIONS RESEARCH*, 499-505.

Gazis, D. C., and Potts, R. B. "The Oversaturated Intersection." *2th International Symposium on Traffic Theory*, London, 221-237.

Gomes, G., and Horowitz, R. (2006). "Optimal Freeway Ramp Metering Using the Asymmetric Cell Transmission Model." *Transportation research. Part C, Emerging technologies*, 14(4), 244-262.

Greenberg, H. (1959). "An Analysis of Traffic Flow." *OPERATIONS RESEARCH*, 7(1), 79-85.

Gunay, B. (2003). "Methods to Quantify the Discipline of Lane-Based-Driving." *Traffic Engineering & Control*, 44(1), 22-27.

Hadi, M., and Wallace, C. (1994). "Optimization of Signal Phasing and Timing Using Cauchy Simulated Annealing." *Transportation Research Record*, 1456, 64-64.

Hadi, M., and Wallace, C. (1995). "Treatment of Saturated Conditions Using Transyt-7f. Ite 65th Annual Meeting." *Compendium of Technical Papers*, 463–466.

Hadi, M. A., and Wallace, C. E. (1993). "Hybrid Genetic Algorithm to Optimize Signal Phasing and Timing." *Transportation Research Record*, 1421, 104.

Hall, R. W. (1993). "Non-Recurrent Congestion: How Big Is the Problem? Are Traveler Information Systems the Solution?" *Transportation Research Part C: Emerging Technologies*, 1(1), 89-103.

Han, B. (1996). "Optimising Traffic Signal Settings for Periods of Time-Varying Demand." *Transportation Research Part A: Policy and Practice*, 30(3), 207-230.

HCM. (2000). *Highway Capacity Manual 2000*, Transportation Research Board, National Research Council, Washtington, D. C.

Henry, J., Farges, J., and Tuffal, J. "The Prodyn Real Time Traffic Algorithm." *Proceedings of the 4th IFAC, Transportation Systems*, 305.

Hou, Z., Xu, J.-X., and Yan, J. (2008). "An Iterative Learning Approach for Density Control of Freeway Traffic Flow Via Ramp Metering." *Transportation Research Part C: Emerging Technologies*, 16(1), 71-97.

Hunt, P., Bretherton, R., Robertson, D., and Royal, M. (1982). "Scoot on-Line Traffic Signal Optimisation Technique." *Traffic Engineering and Control*, 23, 190-2.

Hwang, S. F., and He, R. S. (2006). "Improving Real-Parameter Genetic Algorithm with Simulated Annealing for Engineering Problems." *Advances in Engineering Software*, 37(6), 406-418.

Improta, G., and Cantarella, G. E. (1984). "Control System Design for an Individual Signalized Junction." *Transportation Research Part B: Methodological*, 18(2), 147-167.

Isaksen, L., and Payne, H. (1973). "Suboptimal Control of Linear Systems by Augmentation with Application to Freeway Traffic Regulation." *IEEE Transactions on Automatic Control*, 18(3), 210-219.

ITT Industries, I., Systems Division ATMS R&D and Systems Engineering Program Team. (2006). "Corsim User's Guide." FHWA, US Department of Transportation, Colorado Springs, Co.

Jacobson, L., Henry, K., and Mehyar, O. (1989a). "Real-Time Metering Algorithm for Centralized Control." *Transportation Research Board*, 1232.

Jacobson, L., Henry, K., and Mehyar, O. (1989b). "Real-Time Metering Algorithm for Centralized Control." *Transportation Research Board*, 1232, 17-26.

Jia, B., Jiang, R., and Wu, Q. S. (2004). "Traffic Behavior near an Off Ramp in the Cellular Automaton Traffic Model." *Physical Review E*, 69(5), 56105.

Jong, J.-C., and Schonfeld, P. (2003). "An Evolutionary Model for Simultaneously Optimizing Three-Dimensional Highway Alignments." *Transportation Research Part B: Methodological*, 37(2), 107-128.

Kashani, H., and Saridis, G. (1983). "Intelligent Control for Urban Traffic Systems." *Automatica*, 19(2), 191-197.

Kaya, A. "Computer and Optimization Techniques for Efficient Utilization or Urban Freeway Systems."

Kerner, B. (2005). "Control of Spatiotemporal Congested Traffic Patterns at Highway Bottlenecks." *Physica A: Statistical Mechanics and its Applications*, 355(2-4), 565-601.

Kim, J., Liu, J., Swarnam, P., and Urbanik, T. (1993). "The Areawide Real-Time Traffic Control(Artc) System: A New Traffic Control Concept." *IEEE Transactions on Vehicular Technology*, 42(2), 212-224.

Kirkpatrick, S. (1984). "Optimization by Simulated Annealing: Quantitative Studies." *Journal of Statistical Physics*, 34(5), 975-986.

Kovvali, V., Messer, C., Chaudhary, N., and Chu, C. (2002). "Program for Optimizing Diamond Interchanges in Oversaturated Conditions." *Transportation Research Record: Journal of the Transportation Research Board*, 1811(-1), 166-176.

Lee, S., Messer, C., and Choi, K. (2006). "Actuated Signal Operations of Congested Diamond Interchanges." *Journal of Transportation Engineering*, 132, 790.

Levinson, D., and Zhang, L. (2006). "Ramp Meters on Trial: Evidence from the Twin Cities Metering Holiday." *TRANSPORTATION RESEARCH PART A-POLICY AND PRACTICE*, 40(10), 810-828.

Li, M., and Gan, A. (1999). "Signal Timing Optimization for Oversaturated Networks Using Transyt-7f." *Transportation Research Record: Journal of the Transportation Research Board*, 1683, 118-126.

Li, Z. (2011). "Modeling Arterial Signal Optimization with Enhanced Cell Transmission Formulations." *Journal of Transportation Engineering*, 137, 445.

Li, Z., Chang, G.-L., and Natarajan, S. "Integrated Off-Ramp Control Model for Freeway Traffic Management." *Transportation Research Board 88th Annual Meeting*, Washington D.C., 16.

Lighthill, M., and Whitham, G. (1955). "On Kinematic Waves. Ii. A Theory of Traffic Flow on Long Crowded Roads." *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 317-345.

Lin, W., and Ahanotu, D. (1995). "Validating the Basic Cell Transmission Model on a Single Freeway Link." *PATH Technical Note*, 3.

Lipp, L., Corcoran, L., and Hickman, G. (1991). "Benefits of Central Computer Control for Denver Ramp-Metering System." *Transportation Research Record*, 1320, 3-6.

Little, J., Kelson, M., and Gartner, N. (1981). "Maxband: A Program for Setting Signals on Arteries and Triangular Networks." *Transportation Research Record*, 795, 40-46.

Little, J. D. C. (1966). "The Synchronization of Traffic Signals by Mixed-Integer Linear Programming." *OPERATIONS RESEARCH*, 14(4), 568-594.

Lo, H. K. (1999). "A Novel Traffic Signal Control Formulation." *Transportation Research Part A*, 33(6), 433-448.

Lo, H. K., Chang, E., and Chan, Y. C. (2001). "Dynamic Network Traffic Control." *Transportation Research Part A*, 35(8), 721-744.

Lo, H. K., and Chow, A. H. F. (2004). "Control Strategies for Oversaturated Traffic." *Journal of Transportation Engineering*, 130(4), 466.

Lovell, D., and Daganzo, C. (2000). "Access Control on Networks with Unique Origin–Destination Paths." *Transportation Research Part B*, 34(3), 185-202.

Lovell, D. J. (1997). "Traffic Control on Metered Networks without Route Choice," Ph.D. Dissertation, University of California, Berkeley, USA.

Ma, T., and Abdulhai, B. (2002). "Genetic Algorithm-Based Optimization Approach and Generic Tool for Calibrating Traffic Microscopic Simulation Parameters." *Transportation Research Record: Journal of the Transportation Research Board*, 1800(-1), 6-15.

Masher, D., Ross, D., Wong, P., Tuan, P., Zeidler, H., and Petracek, S. (1975). "Guidelines for Design and Operation of Ramp Control Systems." *NCHRP Project 3-22 Final Report*.

May, A. D. "Friction Concept of Traffic Flow." *Highway Research Board Annual Meeting*, Washington D.C., 493-510.

McCoy, P., Balderson, E., Hsueh, R., and Mohaddes, A. (1983). "Calibration of Transyt Platoon Dispersion Model for Passenger Cars under Low-Friction Traffic Flow Conditions (Abridgment)." *transportation Research Record*, 905, 48-52.

Meldrum, D., and Taylor, C. (1995). "Freeway Traffic Data Prediction Using Artificial Neural Networks and Development of a Fuzzy Logic Ramp Metering Algorithm." *Report No. WA-RD*, 365.

Messer, C., and Berry, D. (1975). "Effects of Design Alternatives on Quality of Service at Signalized Diamond Interchanges." *Transportation Research Record*, 538, 20-31.

Messer, C., Fambro, D., and Richards, S. (1977). "Optimization of Pretimed Signalized Diamond Interchanges."

Michaels, R., and Cozan, L. (1963). "Perceptual and Field Factors Causing Lateral Displacement." *Highway Research Record*, 25, 1-13.

Morgan, J., and Little, J. (1964). "Synchronizing Traffic Signals for Maximal Bandwidth." *OPERATIONS RESEARCH*, 12(6), 896-912.

Munjal, P. (1971). "An Analysis of Diamond Interchange Signalization." *Highway Research Record*(349), 47-64.

Muñoz, L., Sun, X., Sun, D., Gomes, G., and Horowitz, R. "Methodological Calibration of the Cell Transmission Model." 798-803 vol. 1.

Nihan, N., and Berg, D. (1991). "Predictive Algorithm Improvements for a Real-Time Ramp Control System." *Final Report GC8286 (16), WA-RD*, 213.

Paesani, G., Kerr, J., Perovich, P., and Khosravi, E. "System Wide Adaptive Ramp Metering in Southern California." *ITS America 7th Annual Meeting*.

Papageorgiou, M. (1980). "A New Approach to Time-of-Day Control Based on a Dynamic Freeway Traffic Model." *Transportation Research Part B: Methodological*, 14(4), 349-360.

Papageorgiou, M. (1983). "A Hierarchical Control System for Freeway Traffic." *Transportation Research Part B: Methodological*, 17(3), 251-261.

Papageorgiou, M. (1990). "Modelling and Real-Time Control of Traffic Flow on the Southern Park of Boulevard Peripherique in Paris: Part Ii: Coordinated on-Ramp Metering." *TRANSP. RES.*, 24A(5), 361-370.

Papageorgiou, M. (1995). "An Integrated Control Approach for Traffic Corridors." *Transpn. Research. Part C, Emerging Technologies.*, 3(1), 19-30.

Papageorgiou, M., Blosseville, J. M., and Hadi-Salem, H. (1990). "Modelling and Real-Time Control of Traffic Flow on the Southern Part of Boulevard Peripherique in Paris: Part I: Modelling." *TRANSP. RES.*, 24A(5), 345-359.

Papageorgiou, M., Diakaki, C., Dinopoulou, V., Kotsialos, A., and Wang, Y. (2003). "Review of Road Traffic Control Strategies." *Proceedings of the IEEE*, 91(12), 2043-2067.

Papageorgiou, M., Hadj-Salem, H., and Blosseville, J. (1991). "Aline A: A Local Feedback Control Law for on-Ramp Metering."

Papageorgiou, M., and Kotsialos, A. (2002). "Freeway Ramp Metering: An Overview." *IEEE Transactions on Intelligent Transportation Systems*, 3(4), 271-281.

Park, B. B., Messer, C. J., and Urbanik, T. (1999). "Traffic Signal Optimization Program for Oversaturated Conditions: Genetic Algorithm Approach." *Transportation Research Record*, 1683, 133-42.

Park, B. B., Messer, C. J., and Urbanik, T. (2000). "Enhanced Genetic Algorithm for Signal-Timing Optimization of Oversaturated Intersections." *Transportation Research Record: Journal of the Transportation Research Board*, 1727(-1), 32-41.

Payne, H., Brown, D., and Todd, J. (1985). "Demand-Responsive Strategies for Interconnected Freeway Ramp Control Systems. Volume 1: Metering Strategies." FHWA/RD-85/109, VERSAC Incorporated, San Diego, California.

Radwan, A., and Hatton, R. (1990). "Evaluation Tools of Urban Interchange Design and Operation." *Transportation Research Record*(1280), 148-155.

Rahmani, S., Mousavi, S. M., and Kamali, M. J. (2011). "Modeling of Road-Traffic Noise with the Use of Genetic Algorithm." *Applied Soft Computing*, 11(1), 1008-1013.

Reiss, R. A. (1981). "Algorithm Development for Corridor Traffic Control." *Traffic, Transportation, and Urban Planning*, 2.

Richards, P. (1956). "Shock Waves on the Highway." *OPERATIONS RESEARCH*, 42-51.

Robertson, D. (1969). "Tansyt: Method for Area Traffic Control." *Traffic Engineering and Control*, 11(6), 278-281.

Sasaki, T., and Akiyama, T. "Development of Fuzzy Traffic Control System on Urban Expressway." *Control in Transportation Systems*, Vienna, Austria, 215-220.

Schwartz, S., and Tan, H. "Integrated Control of Freeway Entrance Ramps by Threshold Regulation."

Seddon, P. (1972). "Another Look at Platoon Dispersion: 3. The Recurrence Relationship." *Traffic Engineering+ Control*, 13, 442–444.

Sen, S., and Head, K. (1997). "Controlled Optimization of Phases at an Intersection." *TRANSPORTATION SCIENCE*, 31(1), 5-17.

Sims, A., and Dobinson, K. (1980). "The Sydney Coordinated Adaptive Traffic (Scat) System Philosophy and Benefits." *IEEE Transactions on Vehicular Technology*, 29(2), 130-137.

Spall, J. C., and Chin, D. C. (1997). "Traffic-Responsive Signal Timing for System-Wide Traffic Control." *Transportation Research Part C*, 5(3-4), 153-163.

Stamatiadis, C., and Gartner, N. (1996). "Multiband-96: A Program for Variable-Bandwidth Progression Optimization of Multiarterial Traffic Networks." *Transportation Research Record: Journal of the Transportation Research Board*, 1554, 9-17.

Stephanedes, Y. "Implementation of on-Line Zone Control Strategies for Optimal Ramp Metering in the Minneapolis Ring Road." 181.

Stevanovic, A., Martin, P., and Stevanovic, J. (2007). "Vissim-Based Genetic Algorithm Optimization of Signal Timings." *Transportation Research Record: Journal of the Transportation Research Board*, 2035(-1), 59-68.

Stevanovic, J., Stevanovic, A., Martin, P. T., and Bauer, T. (2008). "Stochastic Optimization of Traffic Control and Transit Priority Settings in Vissim." *TRANSPORTATION RESEARCH PART C-EMERGING TECHNOLOGIES*, 16(3), 332-349.

Stringer, S. M. (2006). "Thinking Outside the Box: An Analysis of Manhattan Gridlock and Spillback Enforcement." Office of Manhattan Borough President.

Szu, H., and Hartley, R. (1987). "Fast Simulated Annealing." *Physics Letters A*, 122(3-4), 157-162.

Tabac, D. "A Linear Programming Model of Highway Traffic Control." 568–570.

Taragin, A. "Driver Behavior as Affected by Objects on Highway Shoulders." *Highway Research Board Annual Meeting*, 453-472.

Taylor, C., Meldrum, D., and Jacobson, L. (1998). "Fuzzy Ramp Metering: Design Overview and Simulation Results." *Transportation Research Record: Journal of the Transportation Research Board*, 1634(-1), 10-18.

Tian, Z., Balke, K., Engelbrecht, R., and Rilett, L. (2002). "Integrated Control Strategies for Surface Street and Freeway Systems." *Transportation Research Record: Journal of the Transportation Research Board*, 1811(-1), 92-99.

Tian, Z., Messer, C., and Balke, K. (2004). "Modeling Impact of Ramp Metering Queues on Diamond Interchange Operations." *Transportation Research Record: Journal of the Transportation Research Board*, 1867(-1), 172-182.

Van Aerde, M., and Yagar, S. (1988). "Dynamic Integrated Freeway/Traffic Signal Networks: Problems and Proposed Solutions." *Transportation research. Part A: general*, 22(6), 435-443.

van den Berg, M., De Schutter, B., Hegyi, A., and Hellendoorn, J. (2004). "Model Predictive Control for Mixed Urban and Freeway Networks." Proceedings of the 83rd Annual Meeting of the Transportation Research Board, Washington, DC, Paper 04-3327.

van den Berg, M., Hegyi, A., De Schutter, B., and Hellendoorn, J. "A Macroscopic Traffic Flow Model for Integrated Control of Freeway and Urban Traffic Networks." *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*, 2774-2779 Vol.3.

Venglar, S., Koonce, P., and II, T. (1998). "Passer Iii-98 Application and User's Guide." *Texas Transportation Institute, Texas A&M Univ., College Station, Tex*.

Wang, C. F. (1972). "On a Ramp-Flow Assignment Problem." *TRANSPORTATION SCIENCE*, 6(2), 114-130.

Wang, J., and May, A. "Computer Model for Optimal Freeway on-Ramp Control."

Wattleworth, J. (1963). "Peak-Period Control of a Freeway System–Some Theoretical Considerations." *Chicago area expressway surveillance project*.

Webster, F. (1958). *Traffic Signal Settings*, HMSO.

Webster, F. V., and Cobbe, B. M. (1967). *Traffic Signals*, HMSO.

Wei, C.-H., and Wu, K.-Y. (1996). "Applying an Artificial Neural Network Model to Freeway Ramp Metering Control." *Transportation Planning Journal*, 25(3), 335-356.

Wong, C. K., Wong, S. C., and Lo, H. K. (2007). "Reserve Capacity of a Signal-Controlled Network Considering the Effect of Physical Queuing." *Transportation and Traffic Theory 2007: Papers Selected for Presentation at ISTTT17, a Peer Reviewed Series Since 1959*, 533.

Wong, S. C., and Yang, H. (1997). "Reserve Capacity of a Signal-Controlled Road Network." *Transportation Research Part B*, 31(5), 397-402.

Wu, J., and Chang, G. (1999). "An Integrated Optimal Control and Algorithm for Commuting Corridors." *International Transactions in Operational Research*, 6(1), 39-55.

Xin, W., Michalopoulos, P., Hourdakis, J., and Lau, D. (2004). "Minnesota's New Ramp Control Strategy: Design Overview and Preliminary Assessment." *Transportation Research Record: Journal of the Transportation Research Board*, 1867(-1), 69-79.

Yoshino, T., Sasaki, T., and Hasegawa, T. (1995). "The Traffic-Control System on the Hanshin Expressway." *Interfaces*, 25(1), 94-108.

Yuan, L., and Kreer, J. "Adjustment of Freeway Ramp Metering Rates to Balance Entrance Ramp Queues."

Yuan, L. S., and Kreer, J. B. (1971). "Adjustment of Freeway Ramp Metering Rates to Balance Entrance Ramp Queues." *Transportation Research*, 5(2), 127-&.

Yun, I., and Park, B. B. "Application of Stochastic Optimization Method for an Urban Corridor." *Winter Simulation Conference*, Monterey, California, 1493-1499.

Zhang, H., Ma, J., and Nie, Y. (2009). "Local Synchronization Control Scheme for Congested Interchange Areas in Freeway Corridor." *Transportation Research Record: Journal of the Transportation Research Board*, 2128(-1), 173-183.

Zhang, H., and Ritchie, S. (1997). "Freeway Ramp Metering Using Artificial Neural Networks." *Transportation Research Part C*, 5(5), 273-286.

Zhang, L., and Levinson, D. (2004). "Optimal Freeway Ramp Control without Origin-Destination Information." *Transportation Research Part B: Methodological*, 38(10).

Zhang, M., Kim, T., Nie, X., Jin, W., Chu, L., and Recker, W. (2001). "Evaluation of on-Ramp Control Algorithms." *UC Berkeley: California Partners for Advanced Transit and Highways (PATH). Retrieved from: http://www.escholarship. org/uc/item/83n4g2rq*.

Zhou, G., Gan, A., and Shen, L. (2007). "Optimization of Adaptive Transit Signal Priority Using Parallel Genetic Algorithm." *Tsinghua Science & Technology*, 12(2), 131-140.

Ziliaskopoulos, A. (2000). "A Linear Programming Model for the Single Destination System Optimum Dynamic Traffic Assignment Problem." *TRANSPORTATION SCIENCE*, 34(1), 37-49.