

**Development and Evaluation of Model-Based Adaptive Signal Control  
for Congested Arterial Traffic**

by

Gang Liu

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In

Transportation Engineering

Department of Civil and Environmental Engineering  
University of Alberta

© Gang Liu, 2015

## **Abstract**

Under congested conditions, the road traffic states of different arterial links will interact with each other; therefore, it is necessary to understand the behavior of traffic corridors and to investigate corridor-wide traffic coordinated control strategies. In order to achieve this, traffic flow models are applied in signal control to predict future traffic states. Optimization tools are used to search for the best sequence of future control decisions, based on predictions by traffic flow models. A number of model-based adaptive control strategies have been presented in the literature and have been proved effective in practice. However, most studies have modeled the traffic dynamic either at a link-based level or at an individual movement-based level. Moreover, the efficiency of corridor-wide coordination algorithms for congested large-scale networks still needs to be further improved.

A hierarchical control structure is developed to divide the complex control problem into different control layers: the highest level optimizes the cycle length, the mid layer optimizes the offsets, and the Model Predictive Control (MPC) procedure is implemented in the lowest layer to optimize the split. In addition, there is an extra multi-modal priority control layer to provide priority for different travel modes. Firstly, MPC is applied to optimize the signal timing plans for arterial traffic. The objectives are to increase the throughput. A hybrid urban traffic flow model is proposed to provide relatively accurate predictions of the traffic state dynamic, which is capable of simulating queue evolutions among different lane groups in a

specific link. Secondly, this study expands the dynamic queue concept to the corridor-wide coordination problem. The ideal offset and boundary offsets to avoid spillback and starvation are found based on the shockwave profiles at each signalized intersection. A new multi-objective optimization model based on the preemptive goal programming is proposed to find the optimal offset. Thirdly, the priority control problem is formulated into a multi-objective optimization model, which is solved with a Non-dominated Sorting Genetic Algorithm. Pareto-optimal front results are presented to evaluate the trade-off among different objectives and the most appropriate solution is chosen with high-level information.

Performance of the new adaptive controller is verified with software-in-the-loop simulation. The applied simulation environment contains VISSIM with the ASC/3 module as the simulation environment and the control system as the solver. The simulation test bed includes two arterial corridors in Edmonton, Alberta. The simulation network was well calibrated and validated. The simulation results show that the proposed adaptive control methods outperform actuated control in increasing throughput, decreasing delay, and preventing queue spillback.

## **Acknowledgements**

I would like to take this opportunity to express my deepest gratitude to my supervisor, Dr. Zhi-Jun (Tony) Qiu for his expert guidance, continuous support, and inspiration throughout the course of my study and research. I am greatly influenced by his rigorous attitude toward scientific research and always enjoy discussing research questions with him. Without his timely wisdom and counsel, my dissertation would have been an overwhelming pursuit and never-ending journey.

I would like to thank my fellow graduate students and colleagues at the Centre for Smart Transportation in the University of Alberta which has a wonderful research and collaboration atmosphere. Sincere appreciation goes to Dr. Pengfei Li for his valuable instructions and generous help, and Xu Han for his significant work on the simulation model development. Sincere appreciation also goes to Aalyssa Atley and Rochelle Borchman for providing patient help in technical writing.

Lastly, but definitely not the least, special thanks go to my family for their endless support, sacrifice, encouragement and understanding in the course of all of my studies.

# Table of Contents

<b>ABSTRACT .....</b>	<b>II</b>
<b>ACKNOWLEDGEMENTS.....</b>	<b>IV</b>
<b>TABLE OF CONTENTS.....</b>	<b>V</b>
<b>LIST OF FIGURES .....</b>	<b>IX</b>
<b>LIST OF TABLES .....</b>	<b>XI</b>
<b>LIST OF ABBREVIATIONS.....</b>	<b>XII</b>
<b>CHAPTER 1 INTRODUCTION .....</b>	<b>1</b>
1.1 RESEARCH MOTIVATION .....	1
1.1.1 Characteristics of Congested Arterial.....	1
1.1.2 Traffic Signal Control .....	2
1.1.3 Multi-modal Priority Control .....	4
1.2 STATEMENT OF PROBLEMS .....	5
1.3 RESEARCH OBJECTIVES AND SCOPE .....	7
1.4 RESEARCH CONTRIBUTIONS .....	10
1.5 ORGANIZATION OF THE DISSERTATION .....	12
<b>CHAPTER 2 LITERATURE REVIEW .....</b>	<b>13</b>
2.1 TRAFFIC FLOW MODELS FOR ARTERIAL TRAFFIC NETWORK .....	13
2.1.1 Kinematic Wave Model.....	13
2.1.2 Store-and-Forward Model .....	16

2.1.3 Dispersion-and-Store Model .....	18
2.1.4 Comparison .....	20
2.2 SIGNAL TIMING OPTIMIZATION METHOD .....	22
2.2.1 Simple Prediction based Method.....	22
2.2.2 Advanced Model based Method.....	24
2.2.3 Adaptive Offset Optimiztion Method.....	29
2.2.4 Comparison .....	32
2.3 ON-LINE OPTIMIZATION FRAMEWORK .....	33
2.3.1 Dynamic Programming (DP).....	33
2.3.2 Sequencing Optimization .....	34
2.3.3 Rolling Horizon .....	35
2.3.4 Comparison .....	36
2.4 SUMMARY .....	37
<b>CHAPTER 3 ADAPTIVE SIGNAL CONTROL IMPLEMENTATION AND EVALUATION PLATFORM .....</b>	<b>39</b>
3.1 SOFTWARE-IN-THE-LOOP SIMULATION .....	39
3.1.1 Traffic Controller .....	40
3.1.2 Traffic Microsimulation .....	40
3.1.3 Date Flow and Integration.....	41
3.2 SIMULATION PLATFORM ARCHITECTURE .....	43
3.2.1 ASC/3 Interface .....	44
3.2.2 Control System.....	45

3.2.3 Data Flow and Integration .....	45
3.3 TEST-NETWORK SIMULATION MODEL .....	46
3.3.1 Study Corridors .....	46
3.3.2 Modeling Process .....	49
3.3.3 Calibration and Validation .....	54
3.3.4 Design of Experiments .....	57
<b>CHAPTER 4 PROACTIVE ARTERIAL SIGNAL OPTIMIZATION WITH EMBEDDED ENHANCED STORE-AND-FORWARD MODEL .....</b>	<b>60</b>
4.1 INTRODUCTION .....	60
4.2 ENHANCED STORE-AND-FORWARD MODEL .....	61
4.3 OPTIMIZATION FORMULATION .....	68
4.3.1 Optimization Framework .....	68
4.3.2 Optimization Model .....	70
4.3.3 Solution Algorithm .....	73
4.4 SIMULATION EVALUATION .....	74
4.5 SUMMARY .....	82
<b>CHAPTER 5 ADAPTIVE MODEL-BASED OFFSETS OPTIMIZATION FOR CONGESTED ARTERIAL TRAFFIC .....</b>	<b>84</b>
5.1 INTRODUCTION .....	84
5.2 METHODOLOGY .....	86
5.2.1 Ideal Offset and Boundary Offsets .....	86
5.2.2 Model Formulation .....	92

5.2.3 Solution Algorithm.....	94
5.3 SIMULATION EVALUATION.....	95
5.4 SUMMARY .....	98
<b>CHAPTER 6 ADAPTIVE SIGNAL PRIORITY CONTROL ON MIXED TRAFFIC ARTERIALS.....</b>	<b>99</b>
6.1 INTRODUCTION.....	99
6.2 MODEL DEVELOPMENT .....	102
6.2.1 Problem Formulation.....	103
6.2.2 Solution Algorithm.....	104
6.2.3 Solution Selection.....	107
6.3 SIMULATION EVALUATION.....	109
6.4 SUMMARY .....	118
<b>CHAPTER 7 CONCLUSIONS AND FUTURE WORK.....</b>	<b>119</b>
7.1 CONCLUSIONS .....	119
7.2 LIMITATIONS AND RECOMMENDATIONS.....	121
<b>REFERENCES .....</b>	<b>123</b>

## List of Figures

Figure 1.1 Research Flow Chart.....	10
Figure 2.1 Arterial Link Example of SFM.....	17
Figure 2.2 Concept of Rolling Horizon Scheme [81] .....	36
Figure 3.1 Real Transportation Environment and SILS Data Flow .....	42
Figure 3.2 Framework of Simulation Platform .....	44
Figure 3.3 Data Flowchart.....	46
Figure 3.4 Study Corridors.....	47
Figure 3.5 Modeling Process Flowchart .....	50
Figure 3.6 Examples of Turning Movement at Intersections.....	52
Figure 3.7 Model Calibration Results .....	56
Figure 3.8 Temporal Variations of Traffic Demand .....	58
Figure 4.1 Detector Requirement for Adaptive Systems .....	62
Figure 4.2 Dynamic Traffic Flow Evolutions along Arterial Streets .....	63
Figure 4.3 Framework of the Control Loop .....	69
Figure 4.4 Illustration of the Rolling Horizon Scheme.....	70
Figure 4.5 Typical Vehicular and Pedestrian Movements at a Four-leg Intersection [94] .....	72
Figure 4.6 Standard Ring-and-barrier Diagram [94].....	72
Figure 4.7 GA Process .....	74
Figure 4.8 GA Results for Each Generation.....	76
Figure 4.9 Distribution of Throughput under Different Scenarios.....	79

Figure 4.10 Relative Queue Length Dynamic.....	82
Figure 5.1 Shockwave Profile and Queue Dynamic in Congested Condition .....	87
Figure 5.2 Spillback and Starvation in Congested Condition.....	89
Figure 5.3 Relationships between Offsets for Primary and Opposing Traffic.....	91
Figure 5.4 Average Delay under Different Demand Scenario .....	97
Figure 5.5 Network Reserve Queuing Capacities .....	98
Figure 6.1 Multi-objective Optimization Process .....	102
Figure 6.2 Flowchart of NSGA-II Algorithm .....	106
Figure 6.3 One Example of User Prioritized Rules.....	108
Figure 6.4 Pareto Frontier of Generation 20 from NSGA-II.....	113

## List of Tables

Table 3.1 Signal Timings at Downtown Corridor.....	48
Table 3.2 Signal Timings at Southeast Corridor.....	49
Table 3.3 Turning Movements Data.....	51
Table 3.4 Model Validation Results.....	57
Table 3.5 Intersection Saturation Rate under Different Traffic Demand Conditions	59
Table 4.1 Parameters Setting of GA.....	75
Table 4.2 Throughput Comparison of VISSIM Simulation Results.....	77
Table 5.1 Delay Comparison of VISSIM Simulation Results.....	96
Table 6.1 Summary of Preference-based Adaptive TSP Methods.....	101
Table 6.2 NSGA-II Parameter Used in Simulation Tests.....	109
Table 6.3 Trade-off between Two Objectives.....	114
Table 6.4 Total Bus Travel Times along the Corridor.....	116
Table 6.5 MOEs at Individual Intersections.....	117

## **List of Abbreviations**

---

<b>Acronym</b>	<b>Definition</b>
APC	Automatic Passenger Counters
API	Application Programming Interface
ASC/3	Advanced System Controllers series 3
ATC	Advanced Transportation Controllers
ATSC	Adaptive Traffic Signal Control
ATSP	Adaptive Transit Signal Priority
Ave.	Avenue
AVI	Automatic Vehicle Identification
AVL	Automatic Vehicle Location
GPS	Global Positioning Systems
CTM	Cell Transmission Model
DISCO	Dynamic Intersection Signal Control Optimization
DP	Dynamic Programming
DSM	Dispersion-and-Store Model
EB	East Bound
ETS	Edmonton Transit System
FD	Fundamental Diagram
GA	Genetic Algorithm

---

<b>Acronym</b>	<b>Definition</b>
GPS	Global Positioning System
ITS	Intelligent Transportation Systems
KWM	Kinematic Wave Model
LOS	Level Of Service
LP	Linear Programming
LQ	Linear-Quadratic
LWR	Light hill-Whitham-Richards Model
MILP	Mixed Integer Linear Program
MITROP	Mixed-Integer Traffic Optimization Program
MOE	Measure of Effectiveness
MOTION	Method for the Optimization of Traffic Signals In Online Controlled Networks
MPC	Model Predictive Control
NB	North Bound
NEMA	National Electrical Manufacturers Association
NOC	Nonlinear Optimal Control
NSGA	Non-dominated Sorting Genetic Algorithms
NTCIP	National Transportation Communications for ITS Protocol
OPAC	Optimized Policies for Adaptive Control
PAMSCOD	Platoon-based Arterial Multi-modal Signal Control with Online Data

<b>Acronym</b>	<b>Definition</b>
PI	Performance Index
QPC	Quadratic-Programming Control
RHODES	Real-time Hierarchical Optimized Distributed Effective System
SB	South Bound
SCP	Signal Control and Prioritization
SFM	Store-and-Forward Model
SIGOP	Network Signal Optimization Model
SILS	Software-In-The-Loop Simulation
SQP	Sequential Quadratic Programming
St.	Street
TSP	Transit Signal Priority
TUC	Traffic Responsive Urban Control
UTC	Urban Traffic Control
VISSIM	Traffic in Towns - Simulation (German Acronym)
WB	West Bound

# **Chapter 1 Introduction**

## **1.1 Research Motivation**

### **1.1.1 Characteristics of Congested Arterial**

High levels of traffic congestion during peak periods are regular in busy arterials of major metropolitan areas, because the traffic demand approaches or exceeds the capacity of the arterial network. The identification of congested condition and understanding of the characteristics are the prerequisite to control the congested traffic flow. However, using the precise definition based on demand/capacity ratio is not an easy task in the real world by using the current data collection system. Because it is difficult to measure the actual traffic demand and capacity when the traffic system is congested, the congested condition at signalized intersection can be defined as the condition of having an approach with residual queue [1]. Traffic flow will become unstable under congested conditions. A small fluctuation from any vehicle in a platoon may cause adverse consequences and reduce the efficiency of traffic system sharply. The low stability of saturated traffic flow puts forward more stringent requirements to the traffic control system.

Traffic lights at intersections are the major control measure in urban road network; however, it may lead to less efficient operations when traffic demand approaches or exceeds the network capacity. For example, one limited congestion triggered by a temporarily and locally excessive demand may lead to an unstable

escalation and the creation of secondary congestion when no suitable control actions are employed. Subsequently, it may lead the entire system to restricted mobility and result in degraded operational efficiency [2-4]. If the traffic state comes to the realm of congested condition, traffic intersections are not isolated and the traffic states of roads will interact with each other. Hence, it is necessary to understand the behavior of arterial traffic and to investigate corridor-wide coordinated signal control strategies. Providing an efficient signal control system has become increasingly important because of effects of the high congestion levels on the urban environment and the quality of life. There must exist an optimal control decision to keep the arterial traffic used in a well-organized way.

### **1.1.2 Traffic Signal Control**

Over the past several decades, a large body of literature has been devoted on this vital issue and most of them fall into the following three categories: fixed-time control, actuated control and adaptive control. The fixed-time control strategy in current practice typically segments a day into a number of time intervals, and then a best-suited signal timing plan for each interval is determined by applying Webster's formula or using optimization tools [5-7]. Unfortunately, the anticipated traffic patterns, particularly in congested condition, are seldom realized in the real-world exactly as they were planned. Obviously, the fixed-time control may cause unstable or unreliable control performance. The actuated control strategy partially responds to the real-time traffic arrivals, but it has been proved to be suboptimal control especially under heavy traffic condition because of the preset limits [8, 9]. For

example, it may result in myopic control. Adaptive control strategy adjusts, in real time, signal timing plans in response to real-time traffic flow fluctuations. With advances in computation and sensing, it has become an increasingly attractive option and been researched for the last three decades [10, 11]. Some adaptive control strategies proactively adjust signal timing plans to meet predicted traffic states before vehicles arrive. Others react by providing feedback to the measured traffic states.

Proactive control strategy uses macroscopic, mesoscopic or microscopic traffic flow models to predict the future traffic states, and develop optimization tools to search for the best future control decisions based on the predicted traffic states. Therefore, this strategy, also called as the model-based adaptive control strategy, can make the best control decisions from a long-term point of view. A number of elaborate traffic flow models, which are deductively derived to describe the complex interactions between traffic states evolution and key control parameters, have been applied to provide relatively accurate predictions [12-18]. Subsequently, a number of model-based adaptive control systems have been presented in literature or even implemented in the field, some of which have been proved effective in practice [10]. However, the efficiency of corridor-wide coordination strategies is still needed to be further improved. It is very important to find a trade-off between the accuracy and the computational complexity, so that the model-based control strategies can make better control decisions and also keep being applicable in practice. Furthermore, it remains a challenging task to generate reliable signal timing plans in congested

traffic condition, which can systematically and globally consider the frequently occurred queue interactions among different lanes and adjacent intersections.

### **1.1.3 Multi-modal Priority Control**

Modern arterial traffic flow is usually composed of multiple travel modes (such as cars, transit, pedestrian, trucks, and emergency vehicles), which compete for the same road space [19]. With correct installation and control strategy selection, traffic signals can improve both traffic mobility and safety for all road users. Multi-modal transportation has been recognized as the key to the future sustainable transportation system. Priority control systems are established to favor one mode over another. Transit Signal Priority (TSP) is a popular tool for improving transit performance and reliability, which temporarily adjusts the traffic signal timing to benefit transit vehicles. It is widely accepted that TSP can reduce unintended bus delays at signalized intersections through extending the current green or truncating the current red. A major controversy, though, is that TSP may bring excessive delays on non-TSP approaches, as their assigned greens are shortened [20]. Ever since the emergence of the concept of TSP, researchers and traffic engineers have been seeking for best solutions to implement TSP strategies and improve the level of service for transit operations while offsetting negative impacts to other travel modes [21-23]. Adaptive signal priority control, which combines adaptive signal optimization with TSP strategies, is a cost-effective way to achieve these goals.

In many practical decision-making problems, multiple objectives or multiple criteria are evident. The aforementioned adaptive priority control is a typical multi-

objective optimization problem, where two or more travel mode priority controls must be satisfied simultaneously in order to obtain the preferred signal timing plan. In fact, it is normally the case that priority objectives of different travel modes are in conflict with each other. However, most studies have used the integrated delay as the objective of their priority control algorithms to balance the trade-offs between different travel modes. This is accomplished by combining different objectives through a weighted sum into a single objective, which obviously provides an easy way to enable a decision maker to choose a solution. However, the weighting vector needs to be assumed beforehand in this method. In addition, the weighting factors may not correspond accurately to the relative importance of the objectives.

## **1.2 Statement of Problems**

For different proposed traffic flow models of arterial traffic network, different model-based control strategies have subsequently been derived. Despite the promising results and contribution from previous research, several elements regarding the traffic flow modeling and optimal control problem have not been clearly studied; but they affect the control performance and limit the applicability in congested conditions.

- a. The queue evolution is modeled either at an individual movement-based level or link-based level by most previous studies. Hence, the control optimization problem is impossible to consider the multiple signal phase operation which is usual in practice and the adjusted saturation flow rate when shared lanes exist in a link.

- b. Most previous studies have not explicitly modeled the queue interactions among neighboring lane groups in a link and accounted for the impact of overflow turning queue length, which are very common during congested conditions. For example, left storage bay spillback will occur when turning traffic uses up the entire space of the storage lanes and blocks the through traffic.
- c. Most existing coordination algorithms do not directly consider the dynamic evolution of queues at intersection approaches, so their application to congested conditions may lead to suboptimal results. Instead of using models based on the average traffic flow conditions and predefined congestion conditions, dynamic offset optimization is needed to deal with congestion phenomena.
- d. In the current preference-based approach for multi-modal priority control, a relative preference vector needs to be supplied without any knowledge of the possible consequences. It is obvious that the trade-off solution obtained by using the preference-based method is largely sensitive to the relative preference vector used in forming the composite function.
- e. Most traffic signal controllers in the field do not have the capability to implement adaptive signal control algorithms directly. The simulation platform should be able to test adaptive traffic signal control strategies based on industry standards and protocols. It is much close to the practice

application if the adaptive control strategies can be implemented directly in modern actuated-coordinated traffic controllers.

### **1.3 Research Objectives and Scope**

The overall objective of the research is to develop and evaluate model-based adaptive signal control methodologies for congested arterial traffic flow. This research proposes a hierarchical control structure to divide the signal control problem of a large traffic system into three different control layers. Control problems with different details are addressed in different layers. As shown in Figure 1.1, the highest layer optimizes the cycle length on the basis of flow capacities and volume levels. Over time, the flexible cycle length is updated as the system adapts to changing traffic conditions. The mid layer continuously calculates optimal split with an embedded enhanced SFM and using the rolling horizon scheme for proactive control. The objective is to maximize the throughput and balance the queue length between adjacent links and lane groups. Based on the adjusted cycle length and green splits, the lowest layer adjusts the offsets from the network level by introducing the boundary offsets and considering the spillback offset and starvation offset. In addition, there is an extra multi-modal priority control layer to provide priority to different travel modals in the mixed arterial traffic.

#### **A. Proactive Split Control with Enhanced Arterial Traffic Flow Model**

This component employs traffic flow model to represent traffic state evolutions and their interaction with control parameters over the arterial network. It is an important prerequisite and the goal is to predict the traffic states evolutions into the future. This

research proposes one hybrid traffic flow model for arterial traffic network, which considers shared lanes and the left-turn bay capacity and is capable of simulating the queue evolutions among neighboring lane groups in a link. Then the MPC (Model Predictive Control) approach is adopted to develop a model-based adaptive control strategy which addresses several issues as mentioned in the above section.

### **B. Adaptive Model-based Offset Control**

With embedded traffic flow models for signalized arterial network, traffic state equations and queue evolution equations can be established. Then the coordination problem can be described as a mathematical optimization problem to minimize or maximize a performance measure, which is a complex function of the signal timing, traffic state, and queue dynamic parameters. This research expands the dynamic queue concept to the coordination problem.

### **C. Adaptive Signal Priority Control on Mixed Traffic**

Mixed traffic road users on most urban arterials are controlled by the same set of signals, and must compete for shared road space. Priority control systems are established to favor one traffic mode over another. However, the weighting coefficients need to be assumed beforehand in this method. In addition, the weighting coefficients may not correspond accurately to the relative importance of the objectives or allow trade-offs between the objectives to be expressed. Instead, another way is to use a multi-objective optimization in finding a number of Pareto-optimal solutions. Then, the higher-level information is used to choose one solution.

### **D. Evolutionary Algorithms Development**

Due to the nonlinear nature of the proposed optimization models, an efficient algorithm is needed to provide sufficiently computing efficiency and reliable solutions in real world operations. Many previous studies have shown the effectiveness of Genetic Algorithm (GA) when solving signal optimization problems [24-26]. This study uses GA as the solution algorithms. It starts by initializing a population of solutions. Each individual represents a potential signal timing solution that evolves through many generations. New candidate solutions are generated by crossover and mutation at each iteration [27].

#### **E. Software-in-the-loop Simulation Based Implement and Evaluation**

The proposed adaptive signal control strategies are implemented and evaluated in the software-in-the-loop simulation (SILS) environment. The adaptive control strategies are implemented in the SILS by adding input–output functions over the NTCIP and Transmission Control Protocol/Internet Protocol (TCP/IP). The simulation results verify the performance at the macroscopic level through analytical analysis.

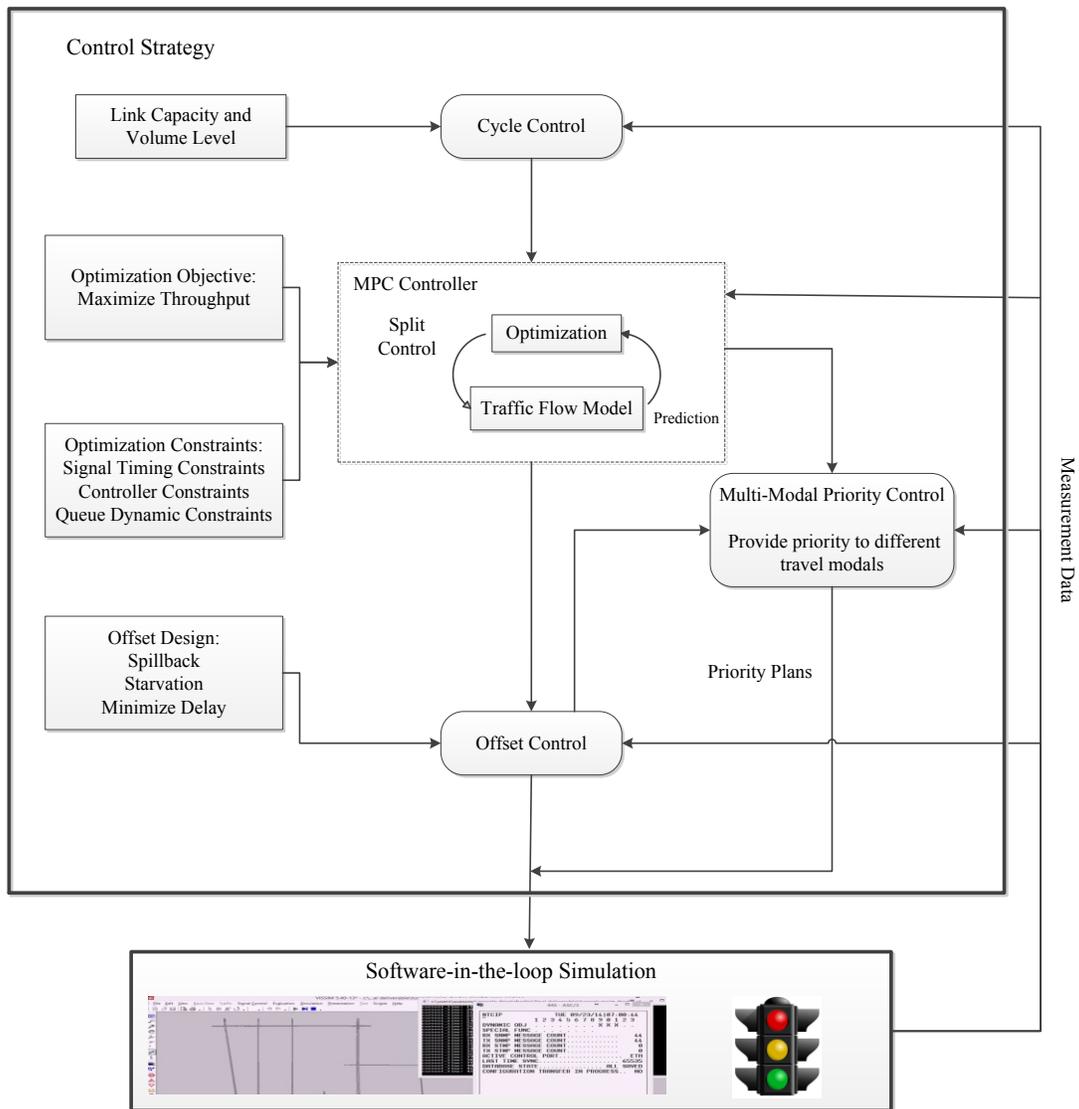


Figure 1.1 Research Flow Chart

### 1.4 Research Contributions

There are several major contributions of this research to the state-of-the-art knowledge in adaptive signal control field, including the followings:

- a. A hybrid lane-group-based traffic flow model is put forward for arterial traffic network by combining the cell-transmission concept, dispersion-and-store concept and store-and-forward concept. It has the potential to offer a

reliable representation of queue evolutions under various types of lane channelization at each intersection approach.

- b. A hierarchical control structure is developed and analyzed for enhanced corridor-wide coordination operations, which aims to maximize the throughput and at the same time prevent the occurrence of starvation, blockage and spillback. The impact of phasing sequence on signal control performance can also be better captured and factored.
- c. A modified rolling horizon scheme is proposed for the successive optimization framework. The control horizon and projection horizon are time-variant, which are based on the implicit timing features of adaptive control and are responsive to the real-time traffic conditions.
- d. An adaptive algorithm is proposed to design signal coordination under congested condition. The ideal offset and boundary offsets to avoid spillback and starvation are found based on the shockwave profiles at each signalized intersection. A new multi-objective optimization model based on the preemptive goal programming is proposed to find the optimal offset.
- e. Instead of using ambiguous weighting factors, multi-objective optimization problems is proposed to generate a set of priority control solutions called Pareto-optimal solution, so that the decision can be taken after the optimization. The most appropriate solution is chosen with high-level information.

## **1.5 Organization of the Dissertation**

There are seven chapters in this dissertation. Chapter 1 gives an introduction of the relevant research background, statement of problems as well as the objectives and scope of this research. The main contributions of this research are also summarized in this chapter. Chapter 2 presents a comprehensive literature review on macroscopic traffic flow models for arterial traffic, signal timing optimization method and on-line optimization framework. Chapter 3 describes the simulation platform architecture, test beds and evaluation scenarios. Chapter 4 presents an enhanced SFM-based signal optimization model to address the queue dynamic and multiple lane groups. Chapter 5 presents an algorithm to design signal coordination for networks with congested intersections. Chapter 6 presents a multi-objective optimization model to find Pareto-optimal front results for evaluating the trade-off among different objectives. Chapter 7 summarizes the main conclusions of this research and discusses recommendations for future research works.

## **Chapter 2 Literature Review**

### **2.1 Traffic Flow Models for Arterial Traffic Network**

A number of macroscopic traffic flow models have been presented in literatures to describe the traffic state evolutions of the urban traffic network. This review focuses only on deterministic models (discretized or not), since only such models could result in the practical formulation of the deterministic optimal control problem for traffic signals. The models, that represent the traffic state evolutions on signalized arterial networks, can be classified into the following three generalized categories: (1) kinematic wave model (KWM); (2) store-and-forward model (SFM); (3) dispersion-and-store model (DSM) [28].

#### **2.1.1 Kinematic Wave Model**

This kind of model is based on the analogies from the hydrodynamic theory. Its general form consists of the two-dimensional conservation equation (Equation 2-1), the definitional formula which states that flow is equal to the product of density and speed (Equation 2-2), and the assumption that the speed is a function of traffic density [29]. In the discretized form, it is assumed that the link is divided into a number of segments. Daganzo proposed the Cell Transmission Model (CTM) which is a convergent numerical approximation to the continuous hydrodynamic model [30, 31]. For a homogeneous roadway, Daganzo suggested using the time-invariant flow-density relationship (Equation 2-3).

$$\frac{\partial k}{\partial t} + \frac{\partial q}{\partial x} = 0 \quad (2-1)$$

$$q = ku \quad (2-2)$$

Where  $k$ ,  $q$  and  $u$  denote traffic density, flow and speed, respectively, which may vary across location  $x$  and time  $t$ .

$$q = \min\{Vk, Q, W(k_{jam} - k)\} \quad (2-3)$$

Where  $k_{jam}$  is the jam density,  $Q$  is the inflow capacity,  $V$  is the free-flow speed, and  $W$  is the backward shockwave speed.

By dividing the whole network into homogeneous cells (cell length equals to the duration of time step multiplied by the free-flow speed), the results of the KWM can be approximated by a set of recursive equations. Equation (2-4) ensures the flow conservation, and Equation (2-5) determines the outflow for each cell at each time step.

$$n_i(t+1) = n_i(t) + y_{i-1}(t) - y_i(t) \quad (2-4)$$

$$y_i(t) = \min\{n_i(t), Q_i(t), \omega[N_{i+1, \max} - n_{i+1}(t)]\} \quad (2-5)$$

Where  $n_i(t)$  = the number of vehicles in cell  $i$  during time step  $t$

$y_i(t)$  = the number of vehicles that leave cell  $i$  during time step  $t$

$N_{i, \max}$  = the maximum number of vehicles that can be accommodated by cell ,

$\omega = W/V$

$Q_i(t)$  = the minimum of capacity flows from cell  $i$  to  $i + 1$

Lo et al. showed that the CTM could be applied to signalized networks. The first cell of one link was modeled to function like a traffic signal and  $Q_i(t)$  was formulated as a binary variable (Equation 2-6) that fluctuated between null and saturation flow rate  $Q_{\max}$  [32-35]. In Lin and Wang's model [36], cells in the network were categorized into four groups: ordinary, intersection, origin, and destination. However, the two models were only applicable to one-way traffic.

$$Q_i(t) = \begin{cases} Q_{\max} & \text{If } t \in \text{green phase} \\ 0 & \text{If } t \in \text{red phase} \end{cases} \quad (2-6)$$

Zhang et al extended the two above models to two-way traffic and all the cells composing the network were categorized into five groups: ordinary, origin, destination, non-signalized diverge, signalized diverge, and signalized merge cells [26]. The origin cells were those with the inflow fixed as the corresponding demand input, and the destination cells were those with outflow unlimited. Li proposed enhanced CTM formulations to consider queue blockage among different lane groups at an intersection approach [37]. Each link was conceptually divided into four zones: the merging, propagation, diverging, and departure zones. Vehicles entering such a link moved over these four zones and then proceed to their respective destinations. Because in the diverging zone vehicles bounded to different destinations could join different queues, the enhanced CTM could consider blockage among different movements.

### 2.1.2 Store-and-Forward Model

Store-and-forward model was first proposed by Gazis for representing the traffic conditions at oversaturated intersections and had since been used in various works notably for road traffic control [38, 39]. The concept was essentially adopted from the theory of communication networks. In this modeling approach, it was first assumed that vehicles entering a link were traveling at a fixed travel time. Then the vehicles were either stored at the end of this link in case of red signal, or further forwarded to downstream links at saturation flow rate during the time of green [28].

Considering a link  $z$  connecting two intersections  $i-1$  and  $i$  (Figure 2-1), the traffic dynamic of link  $z$  is given by the conservation equation (Equation 2-7) [16]. Queues are subject to the Equation 2-8. During periods of high demand, this constraint may automatically lead to a suitable upstream gating for protecting downstream areas from oversaturation. The inflow to the link  $z$  is given by Equation 2-9.

$$x_z(k+1) = x_z(k) + T[q_z(k) - s_z(k) + d_z(k) - u_z(k)] \quad (2-7)$$

$$0 \leq x_z(k) \leq x_{z,\max} \quad (2-8)$$

$$q_z(k) = \sum t_{i,z} u_i(k) \quad (2-9)$$

Where  $T$  = the discrete-time step

$k = 0, 1, \dots$ , the discrete-time index

$x_z(k)$  = the number of vehicles within link  $z$  at time  $kT$

$q_z(k)$ ,  $u_z(k)$  = the inflow and outflow of link  $z$  in the period  $[kT, (k+1)T]$ ,

respectively

$d_z(k)$ ,  $s_z(k)$  = the demand and the exit flow within the link, respectively

$x_{z,\max}$  = the maximum admissible queue length, number of vehicles

$t_{i,z}$  = the turning ratio towards link z from the links that enter intersection  $i$

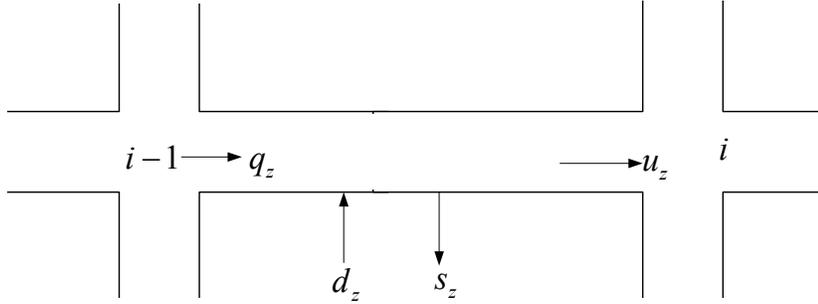


Figure 2.1 Arterial Link Example of SFM

The most important characteristic of SFM is that the discrete-time step  $T$  is equal to cycle length, which enables the mathematical description of the traffic flow evolution without use of discrete variables. The outflow  $u_z(k)$  then has an average value for each period (Equation 2-10). This is of paramount importance because it opens the way to the application of a number of highly efficient optimization and control methods with polynomial complexity, which allows for coordinated control of large-scale networks in real time [16, 40].

$$u_z(k) = G_z(k) \cdot S_z / C \quad (2-10)$$

Where  $G_z(k)$  is the green time of link  $z$ ;  $S_z$  is the saturation flow rate of link  $z$ ; and  $C$  is the cycle length.

The SFM is a simple model and is only applicable in the congested condition, when the vehicle queues resulting from the red phase cannot be dissolved completely at the end of the following green phase. Later, the model was extended to represent

of all possible traffic conditions (congested as well as uncongested) [13, 41]. In these approaches, a nonlinear outflow function was defined (Equation 2-11). However, a continuous link outflow (rather than zero flow during red and free flow during green), was still maintained.

$$u_z(k) = \min \{G_z(k) \cdot S_z / C, x_z(k) / T\} \quad (2-11)$$

### 2.1.3 Dispersion-and-Store Model

Dispersion-and-store model is based on empirical observations to simulate the dispersion of a platoon, that is, platoon of vehicles entering a link are dispersed until they are uniformly distributed on the link stretch. The dispersed platoon is subsequently either stored at the end of the link when the signal turns to red, or further diffused on the downstream link when the signal stays in green [42]. A number of literatures have developed different models to describe the behavior of platoons between signalized intersections. Generally, there are two kinds of mathematical models describing the dispersion of a platoon: Normal Distribution Model proposed by Pacey and Geometric Distribution Model proposed by Robertson [43-46].

Research had already been conducted on the applicability of platoon dispersion model as a reliable traffic flow model in urban networks. Most of the research had shown that Robertson platoon dispersion model is reliable, accurate, and robust [46-48]. It has become a virtually universal standard for platoon dispersion model and has been implemented in some traffic simulation software. The

basic Robertson platoon dispersion model takes the following mathematical form [49].

$$q_t^d = F_n * q_{t-T} + (1 - F_n) * q_{t-n}^d \quad (2-12)$$

Where  $q_t^d$  = the arrival flow rate at the downstream signal at time t

$T_a$  = the average link travel time

$T$  = the minimum travel time on the link (measured in terms of unit steps

$T = \beta T_a$ )

$q_{t-T}$  = the departure flow rate at the upstream signal at time t-T

$n$  = the modeling time step duration

$F_n$  = the smoothing factor given by

$$F_n = \frac{1}{1 + \alpha_n \beta_n T_a} \quad (2-13)$$

Where  $\alpha_n$  is the platoon dispersion factor and  $\beta_n$  is the travel time factor.

Equation (2-12) shows that the traffic flow  $q_t^d$  is a weighted combination of the arrival pattern at the downstream end of the link during the previous time step  $q_{t-n}^d$  and the departure pattern from the upstream traffic signal T seconds ago  $q_{t-T}$ . As it is an empirical model, the accurate calibration is critical in developing effective traffic signal timing plans. The state of practice has been the use of a goodness-of-fit approach to calibrate the model parameters. Alternatively, Yu developed an analytical framework for calibrating parameters of the platoon dispersion model using a statistical analysis of the link travel time distribution [50]. Rakha and Farzaneh improved Yu's procedure and developed three generalized platoon

dispersion models that explicitly accounted for the effect of the time step duration on platoon dispersion [51]. Wong et al. modified the original DSM to deal with the problem of time-varying demand. It employed a calibrated set of sheared formulae for queues and delays, which was based on the group-based signal specification [52].

#### **2.1.4 Comparison**

CTM is capable of describing the traffic flow phenomena under the entire span of traffic conditions. It calls for the subdivision of network links into shorter cells and correspondingly shorter time steps. Thus CTM describes the link-internal traffic state evolutions more accurately. For example, it can capture shockwaves and queue dynamic phenomena. Unfortunately, it has the following obvious disadvantages: (1) the real-life implementation of CTM-based optimization control faces some difficulties, because the creation of large dimensional state vectors results in high computational requirements; (2) the real-time application calls for specific measurements for each cell which are usually not available or highly noisy due to various effects. It seems CTM has a limited significance in interrupted (signal-controlled) traffic flow, in contrast to the uninterrupted freeway traffic flow, because many unpredictable and hardly measurable disturbances (incidents, illegal parking, pedestrian crossings, intersection blocking, etc.) may perturb the traffic flow in urban areas [16].

Clearly, SFM is a simple model and it can only provide a rough representation of the traffic dynamic in congested condition. For example, the model is not aware of short-term queue oscillations due to green-red switching within a

cycle. However, it represents the stop-and-go traffic flow dynamic of signalized arterial network fairly well in heavy congested condition, because the uncongested part of a link is considered negligible compared to the total link length and a platoon cannot be dispersed. Averaging link outflow over one or more cycles, which is determined as a percentage (green per cycle ratio) of the saturation flow rate, implies that it is only applicable for split optimization. It is obvious that cycle length and offsets have no impact within the SFM and must be either fixed or updated in real-time independently. Finally, the linear state-space feature of the store-and-forward model opens the way to the application of a number of highly efficient optimization and control methods.

Although DSM is only an empirical model, it is generally considered to represent interrupted traffic flow in signalized networks better in moderate traffic condition. This model has been empirically validated in several urban areas around the world. It is also known that in heavy congested condition the CTM model predicts a complicated queue evolution, where queues could be formed and dissipated at various locations along a link. However, the real-time accurate calibration of the model parameters is difficult.

On the whole, both SFM and DSM consider the whole link as a single storage segment, but CTM uses the discretization of time and space in order for the continuous model to be approximated by a set of finite difference equations.

## **2.2 Signal Timing Optimization Method**

### **2.2.1 Simple Prediction based Method**

In the 1980s and 1990s, a number of model-based adaptive control systems emerged, such as OPAC in USA [53, 54], PRODYN and CRONOS in France [55, 56], MOTION in Germany [57] and UTOPIA in Italy [58]. Several experimental studies have shown benefits obtained by these systems on the delay and travel time compared to actuated signal timing plans.

The prediction methods of these systems are similar, which predict the future traffic arrivals through the historical data measured from the upstream detectors or the detectors of upstream links. For example, the ideal detector location for OPAC is about 10s upstream of the stop-line (at free flow speed) or upstream of the worst queue on each lane of all through phases [10]. PRODYN estimates the number of vehicles lying between a detector and the traffic signal by using a upstream detector [55]. This kind of prediction methods can take into account the traffic flow variations at a scale of a few seconds and more globally (at the level of the intersection) than the actuated control. However, it is obvious that they are limited in the length of the time horizon. The longest prediction horizon is the time taken by the vehicles running from the upstream detector to the stop-line of the intersection.

These systems use optimization methods to determine the green phase duration by a small time steps (4 or 5 s). The cycle duration is not constrained and varies from one cycle to the next. Only a few systems adjust or optimize phase sequence [59]. The obvious advantage is the greater flexibility for finding the green

phase durations in response to the real-time traffic states, especially for those which have a wide possible cycle spectrum at each cycle. However, they are not able to optimize several intersections of a large-scale network in the same optimization process, because most of the used optimization methods behave exponentially with the number of intersections. Some of these optimizations use heuristic techniques and extensive search techniques to find solutions. For example, OPAC employs complete enumeration while PRODYN employs dynamic programming and decision tree [55, 60, 61]. On the other hand, UTOPIA employs a heuristic global optimization method with polynomial complexity which allows for simultaneous consideration of several intersections [58]. This difficulty leads to a sub-optimality control for a large-scale network.

Most of the systems operate on two or more hierarchical levels, which divide the complex control problem of a large traffic system into different control levels or layers. The lower control level mainly focuses on local control in a more elaborate way, and the higher control level deals with network-wide coordinated control in a more general way [60]. For example, the RHODES system is developed into a three-level hierarchical structure [61, 62]. At the highest level, there is a dynamic network loading model that captures the slow-varying characteristics of traffic, such as road closures and construction. At the middle level, network flow control is actuated to coordinate road network, which is based on the prediction and estimation of the traffic flow loads on the roads. At the bottom level, intersection control is carried out by applying a model-based rolling horizon optimization approach.

## **2.2.2 Advanced Model based Method**

### **A. CTM based Methods**

Dynamic Intersection Signal Control Optimization (DISCO) was the first dynamic urban traffic optimization control approach based on CTM [15, 33, 34]. As discussed in section 2.1, by introducing binary variables, equation 2-5 was equivalently converted into a linear system. DISCO considered the entire fundamental diagram of traffic flow, which was essential for controlling congested and transient traffic. DISCO was able to generate a dynamic timing plan and optimized cycle length, phase splits, and offsets explicitly. The timing plans were derived by solving the optimization problem via a genetic algorithm. DISCO was applied to a congested network in Hong Kong and the results showed that DISCO outperformed the existing plans by 30–40% in overall delay reductions [35].

In order to handle the number of stops, fixed or dynamic cycle length and the problem of unintended vehicle holding, Lin and Wang proposed an enhanced 0-1 mixed-integer linear programming formulation based on CTM, in order to minimize a weighted sum of total delay and total number of stops [36]. Equation 2-5 was replaced by three linear inequalities that did not accurately replicate flow propagation and might suffer the so-called “vehicle-holding problem.” To address this issue, a penalty term for the phase change was used in the objective function to capture the cost associated with the lost time. The model was capable of capturing physical queues, fixed and variable cycle length, and the number of stops, while preserving the minimum and maximum green durations. In terms of computational

efficiency, the proposed formulation had the least number of binary integers as compared with other previous formulations that were developed with the same approach.

Pohlmann et al. developed a prototype of a new adaptive control based on CTM [63]. Every 15 minutes the proposed prototype adapted and optimized signal plans and coordination patterns to the currently estimated traffic demand in the network. Firstly, the upcoming traffic demand was forecasted and estimated. Then cycle length and green splits were adjusted based on the estimated demand. Finally, offsets were optimized by using the CTM in combination with Genetic Algorithms and a second alternative approach named Sequential Enumeration. The simulation results showed that quality of the optimized signal plans and especially the adequacy of cycle lengths stood and fell with the accuracy of traffic demand.

Li presented an arterial signal optimization model to captures traffic dynamic with enhanced CTM formulations, which took into account complex flow interactions among different lane groups [37]. The signal optimization model could optimize the cycle length, split, and offset, while preventing link blockage and lane blockage. Extensive simulation experiments were conducted for field segments of four congested intersections in Silver Spring, Maryland. Through comparisons with signal-timing plans from TRANSYT-7F, results demonstrated that both the total delay and throughput resulting from the proposed model were far better, particularly in congested condition.

## **B. SFM based Methods**

Presumably the first report on the use of SFM for signal control optimization is by Dans and Gazis [64]. They formulated the problem of minimization of the aggregate delay as a dynamic optimization model. Through time discretization, the model was reduced to a linear programming (LP) problem for a fairly wide range of operational conditions, in order to obtain the order of queues exhaustion and an approximation to their optimum time variation. A variety of constraints, such as an upper limit on the individual delay, were transformed into additional linear constraints of the LP problem. The method is demonstrated by optimizing the operation of a two-node, four-queue network.

Later, Diakaki developed the TUC (traffic responsive urban control) system by using SFM as the underlying traffic flow model [65-67]. Instead of optimizing the signal timing parameters, TUC optimized the linear multivariable feedback regulator off-line. The control law was developed through the application of the Linear-Quadratic (LQ) methodology to the formulated optimal control problem. It required the availability of nominal values of green splits, which were the values of green splits that were optimal for a given historical demand and might be obtained through available techniques (e.g. through TRANSYT optimization).

Compared with TUC, Aboudolas et al. presented other two novel control methodologies based on the SFM [16]. Firstly, an open-loop quadratic-programming control (QPC) approach was developed, which can be efficiently solved by using broadly available codes of commercial software. However, to keep the linear characteristic, the store-and-forward model was only applicable in congested

condition. Therefore, an open-loop nonlinear optimal control (NOC) approach was developed based on a nonlinear traffic model, which was more elaborate to describe more complex traffic dynamic. A numerical feasible-direction optimization algorithm was applied to solve NOC iteratively, which required more computational complexity than QPC. A preliminary simulation-based investigation was conducted to demonstrate the comparative efficiency and real-time feasibility of the developed signal control methods.

Later, Aboudolas et al. investigated the efficiency of the QPC that aimed at balancing the link queues and minimizing the risk of queue spillback [40]. The corresponding optimization algorithm was embedded in a rolling-horizon control scheme for the application of the proposed methodology in real time. The efficiency and real-time feasibility was demonstrated and compared with the LQ approach via simulation test with a number of different demand scenarios.

### **C. DSM based Methods**

TRANSYT is one of the most widely used offline signal optimization programs and uses the DSM as the traffic flow model. The link flows and link turning proportions are inputs, which are assumed to be constant for the entire simulation period. The performance index is a combination of the total delay and the number of stops made by vehicles. If the adjusted timings improve the performance index, the optimization process will output the beneficial timings [68]. The optimum is reached by successive adoption of beneficial timings. Its hill-climbing optimization algorithm does not guarantee that a global optimum will be achieved and is also highly

dependent on the quality of the starting solution. Another limitation of TRANSYT is that the performance is questionable in heavy congested traffic condition. Version 7 of TRNSYAT was modified by Federal Highway Administration in 1981 to accommodate driving on the right. TRANSYT-7F incorporates genetic algorithm search technique to improve and accelerate the convergence to optimal solution. In the recent releases of TRANSYT-7F, new objective functions are added to handle heavy congested condition (i.e., minimize queue and maximize throughput) [69].

SCOOT was first developed by Robertson's team and has been extended later in several respects. It is has been applied to over 150 cities in the world [10]. SCOOT incorporates an optimizer into the TRANSYT for online application and includes algorithms for dynamic control of individual intersections, arterials, and grids/networks. Similar to TRANSYT, SCOOT seeks to minimize the linear combination of vehicular delay and stops. It uses link flow profile to tune cycle length, splits, and offset values of each intersection on cycle-by-cycle basis. More precisely, SCOOT is run repeatedly in real time to investigate the effect of incremental changes of splits, offsets, and cycle time. If the changes turn out to be beneficial, they are submitted to the local signal controllers. SCOOT handles congestion with several features, such as: congestion importance factors, congestion offset, gating, and variable node-based target saturation. However, if queuing occurred right up to the exit detector, SCOOT is not capable to model this condition and could not detect the stationary vehicles [70].

### **2.2.3 Adaptive Offset Optimization Method**

The literature review on the adaptive offset optimization methods, which can be classified as two types: 1) the centralized control, and 2) the hierarchical control.

#### **A. Centralized Control**

Wey and Jayakrishnan [71] presented an integer-linear program of signal optimization with an embedded Robertson's platoon dispersion model. The model assumed flexible cycle lengths and phase sequences and included explicit constraints to model the movement of traffic along the streets, and to capture the permitted movements from signal controllers. Lo et al. presented the Dynamic Intersection Signal Control Optimization (DISCO) prototype, which may be the first dynamic urban traffic optimization control approach based on the Cell Transmission Model CTM [35]. DISCO was able to generate a dynamic timing plan and optimized cycle length, phase splits, and offsets explicitly. Later, several other studies extended and improved the CTM-based signal timing optimization [26, 36]. Li extended the cell transmission concept to take into account complex flow interactions among different lane groups. The proposed arterial signal optimization model can yield effective signal plans for both saturated and under-saturated intersections [72].

MITROP (Mixed-Integer Traffic Optimization Program) was designed to simultaneously optimize all the traffic control variables of the network including cycle time, splits of green time, and offsets. The traffic flow dynamic was described both by deterministic and stochastic models. The optimization problem was formulated in terms of mixed-integer linear programming and a globally-optimal

solution was determined using IBM's MPSX optimization system [73]. The method was applied to several traffic signal networks.

The work by Abu-Lebdeh and Benekohal [74, 75] provided frameworks for developing a signal coordination model on arterials with oversaturated intersections. The works were based on the dynamic queue management of a signal system on a single arterial. The split, cycle length and offsets were dynamically and continuously adjusted to respond to real-time conditions. Girianna and F. Benekohal extended the concept of signal coordination to a grid network of oversaturated arterials, and formulated the signal coordination as a dynamic optimization problem. The algorithm intelligently generated optimal signal timing plans along individual arterials by considering the traffic demand's variation and the position of critical signals [76].

Recently, Liu and Chang proposed an optimization model for the design of arterial signal timings with an embedded set of enhanced macroscopic traffic flow equations, which can precisely model the traffic evolution along the arterial link [77]. He et al. presented a unified platoon-based mathematical formulation, called PAMSCOD, to perform arterial traffic signal control [17]. A mixed-integer linear program (MILP) was solved to determine future optimal signal plans (cycle length, offset, split) based on the current traffic controller status, online platoon data and priority requests from special vehicles.

## **B. Hierarchical Control**

SIGOP (Network Signal Optimization Model) consisted of two major components: a flow model and an optimization methodology. The objective function was expressed as system disutility in terms of vehicle delay, stops, and excess queue length. The optimization procedure sought the optimal signal setting to minimize the value of disutility [78].

TRANSYT was one of the most widely used offline signal optimization programs and uses the platoon dispersion flow model. The link flows and link turning proportions were inputs, which were assumed to be constant for the entire simulation period. The performance index was a combination of the total delay and the number of stops made by vehicles. If the adjusted timings improved the performance index, then the optimization process outputs beneficial timings. Optimization is reached by the successive adoption of beneficial timings [68].

Lieberman et al. proposed the RT/IMPOST [79]. The idea was to control queue growth on every saturated approach by suitably metering traffic to maintain stable queues. A mixed-integer linear program (MILP) was formulated to yield optimal values of signal offsets and queue length for each approach. In order to continuously control the actual queue lengths on each saturated approach at optimal queue lengths computed by the MILP formulation, a nonlinear programming formulation adjusted the green phase durations of each signal cycle.

Diakaki et al. developed the traffic-responsive urban control (TUC) to provide coordinated, traffic-responsive control in large-scale urban networks [67]. It included four parts: split control, offset control, cycle length control and public

transport priority. A decentralized feedback control law was applied to effectuate the offset control. Taking into account the possible existence of vehicle queues, it modified the offsets of the main stages of successive junctions along arterials to create green waves.

Based on the CTM, Von der Fakultät für proposed a new offset optimization method for signalized arterial networks. The method consisted of three modules: (1) the input module; (2) the optimization module consisting of a Genetic Algorithm (GA) based optimizer; and (3) a traffic analysis module that serves as the fitness function for the GA-based optimizer [80]. In Pohlmann and Friedrich's research, every 15 minutes the ATCS adapted and optimized signal plans and coordination patterns to the current estimated network traffic demand. In the first step, the upcoming traffic demand was forecasted and estimated. Based on this demand, cycle length and green splits were adjusted. Finally, offsets were optimized by using the CTM in combination with GA and a second approach, called Sequential Enumeration [63].

#### **2.2.4 Comparison**

By employing traffic flow models fed with traffic measurements, the corresponding signal control problem is readily formulated to a dynamic optimization problem. It usually includes discrete variables to reflect the impact of red/green phases on traffic flow. Several constraints, such as maximum and minimum splits, are included. From the literature, we find the real-time solution and realization faces a number of apparently difficulties. This is probably why heuristic solution algorithms are

devised in order to solve the optimization problem. Indeed, the heuristic algorithms reduce the solution time of the problem.

The reviewed systems adjust three major types of signal timings: green splits, cycle length, and offsets; however, it seems the existing research do not study the following issues enough. Firstly, little research fully studied the impact of phase sequence optimization on control performance. Especially, if the traffic flow model captured queue interactions in a link, the optimization model would factor the impact of phase sequence easily. Secondly, most existing signal coordination algorithms for congested networks do not directly consider the dynamic evolution of queues. Most reported adaptive control systems are unable to find the exact optimal offsets. Thirdly, Webster's formula for calculating cycle length is invalid when saturation level exceeds 1.0. Little research proposes a clear method to optimize the cycle length in congested condition. Finally, the reviewed control strategies seems outmoded as compared with the real-life signal strategies that assume dual-ring, 8-phase, variable cycle and phase controllers.

## **2.3 On-line Optimization Framework**

### **2.3.1 Dynamic Programming (DP)**

Dynamic programming is an exact solution for optimization over time. It decomposes a control problem to a series of sub-problems (i.e. step), which corresponds to discrete segments of time in real-time control problem. At each step, a set of state variables give the information on the controller and the traffic states at that time. The Bellman's equation is recursively calculated backwards step-by-step

to find the optimal action, which transfers the system from the current state to a new state. In summary, the DP is a global optimization strategy for multistage decision processes and it provides a standard with which all other strategies can be compared. Application of DP to the signal control problem can be found in [53, 81].

Unfortunately, the implication of DP for real-time traffic signal control is limited. Firstly, the computational demand is exponential to the size of the state space, the information space and the action space. Furthermore, in practice it is difficult to obtain the complete information on the time period in which the controller seeks optimization. For example, traffic detectors may supply only 5–10s data of future arriving vehicles. Finally, most of the outputs from the program are never implemented because optimized policies are generated for all possible combinations of initial conditions at each stage of the control period. In practice, only one optimum policy would be implemented. By being able to produce the theoretically optimal control strategy for each input state, DP usually serves as a standard for evaluation of the relative effectiveness of other strategies that can be implemented in practice.

### **2.3.2 Sequencing Optimization**

Sequencing optimization has the following features: (1) the control period is divided into stages; (2) each stage is divided into an integral number of intervals; and (3) during each stage there must be a sufficient number of phases to guarantee that no optimal solution is missed. The phase-change (switching) times are measured from the start of the stage. Then the optimization problem is to determine the sequence of

switching times to minimize the delay over the whole state. At each state, the initial queues on each approach and the arrivals of the stage are given. The solution of the problem is to search the set of all possible combinations of valid switching times within the stage to determine the optimum sequence [60].

Although sequencing optimization procedure needs the information of vehicles arrivals over the entire stage length, it is more readily to operation in real-time. Obtaining accurate arrivals over this length of time is difficult, but it could be implemented with a traffic prediction model that predicts the traffic pattern over the entire stage.

### **2.3.3 Rolling Horizon**

In this scheme, a projection horizon is predetermined which consists of  $N$  time intervals, as shown in Figure 2.2 [82]. Traffic states are measured for the first  $H$  intervals (head portion) and are estimated from the traffic flow model for the next  $N-H$  intervals (tail portion). Optimal control actions for the whole projection horizon are specified based on the measured and predicted information so as to optimize the performance indices over a target period. However, each control step only implements the first control sample of the optimal control sequence. Then, the projection horizon is shifted into the future by one roll period and the same process is restarted again with new traffic measurements for the next iteration. Usually, the roll period is equal to the length of the head portion.

Because the rolling horizon scheme considers a much longer future period for signal optimizations, it can avoid myopic decisions and achieve better signal

operations. Furthermore, it can coordinate multiple control measures and objectives, and is more robust to disturbances and model mismatch errors.

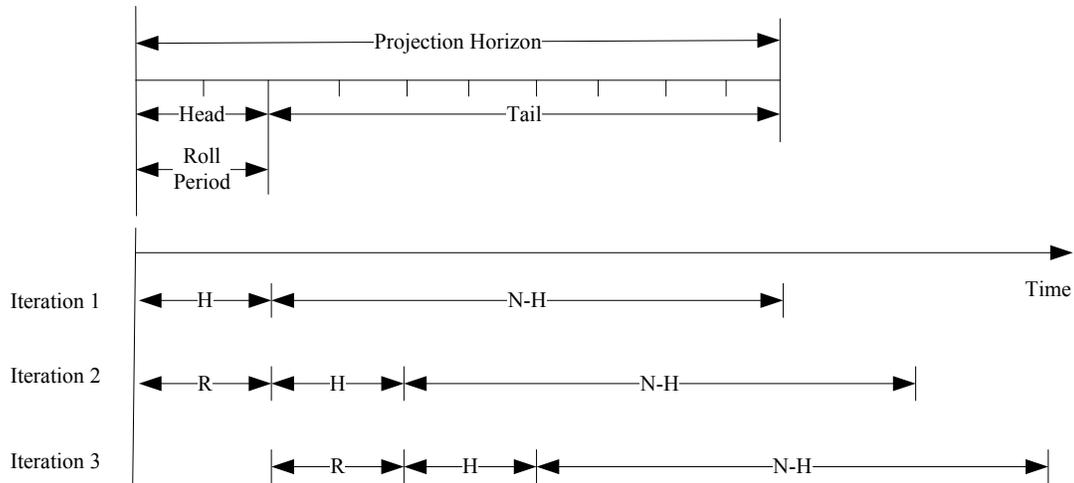


Figure 2.2 Concept of Rolling Horizon Scheme [81]

### 2.3.4 Comparison

Most of the reviewed adaptive control systems adopt the rolling horizon procedure. Hereby, the optimization problem is solved over a projection horizon  $N$ , by using measured initial traffic measurements and demand predictions over  $N$ . After new measurements are collected and a new optimization problem is solved, and so forth. In the practice of signal control, the previous research has several disadvantages. First, most research assumes the length of the projection horizon is pre-set and fixed. If the traffic demand is relatively low and stable, the longer horizon is unnecessary. Because the traffic states for the tail portion are predicted from the traffic flow model, the efficiency of rolling horizon approach may be seriously affected if the prediction is inaccurate within such a long tail portion. Secondly, using short intervals, such as

in DISCO, places a heavy burden on computational requirements and also leads to operational inefficiency.

## **2.4 Summary**

In summary, model-based adaptive control strategies have been developed for a long period of time, and the results are fruitful. A number of adaptive control systems have been presented in literatures or even applied in practice. Some of these systems, which were implemented in real-life traffic field, have been proved effective in practice. However, the efficiency of corridor-wide strategies for large-scale traffic networks is still needed to be further improved.

A number of elaborate urban traffic models, which are deductively derived to describe the traffic flow dynamic, have been applied. For different traffic flow models, different model-based control strategies have subsequently been derived. Unfortunately, the development of corridor-wide model-based control strategies faces obvious difficulties due to the combinatorial nature of the related optimization problem. For example, if the traffic flow models are nonlinear, the computational complexity will increase exponentially when the scale of the network grows. As a consequence, any feasible strategy design includes some simplification, either in its traffic flow model, or in its optimization model and algorithm, or in the control structures. Generally speaking, it is very important to find a trade-off between the accuracy and the computational complexity of the model, so that the model-based control can make better control decisions and also keep being applicable in real-life practice.

Traffic flow models help adaptive control systems perform more proactively, although they also may introduce errors that can be propagated (spatially and temporally) during the course of control actions. MPC is a methodology that implements and repeats optimal control in a rolling horizon scheme. It is a closed-loop control by integrating the real-time feedback. Hence, it is able to deal with the unpredictable disturbances, traffic demand variation and mismatch errors of the prediction model.

## **Chapter 3 Adaptive Signal Control Implementation and Evaluation Platform**

### **3.1 Software-in-the-Loop Simulation**

Simulation platform is needed to implement and evaluate Adaptive traffic signal control (ATSC) strategies based on the following reasons. Firstly, modern traffic controllers have specific physical architectures, control logic, data flows, communication interfaces, and protocols. Vendor-specific controller capabilities are typically do not support the ATSC strategies directly. It is hard to guarantee the transferability of ATSC strategies from the research to the field. Therefore, testing ATSC strategies on actual traffic controllers prior to field implementation is necessary to bridge this gap. Secondly, a field test with adaptive signal controller requires detector installation, backhaul communication setup and other maintenance activities, which are expensive. Many ATSC strategies have been tested using microsimulation packages, which can simulate the signal control and microscopic drives behaviors on urban arterials. The latest advance is a concept called Software-in-the-loop Simulation (SILS), which retain the functionality of a real-world traffic controller. It consists of a microscopic simulation model and several virtual traffic controllers under the simulation software. The communication and exchange of information between these two components are achieved by a controller interface.

For example, Econolite's Advanced System Controller series 3 (ASC/3) is linked to VISSIM [83]. This study uses the ASC/3 SIL controller embedded in VISSIM.

### **3.1.1 Traffic Controller**

Functions of ASC/3 controller include control, coordination, preemption and TSP features, extent detector options, and communication abilities [84]. The virtual ASC/3 controller in the SILS performs identically as hardware controller, and they runs from the same code base. Complex signal timing plans can be realized by the logic processor, where different commands can be either accessed directly or enabled through a special extension file. The emulated external logic provides the capabilities to implement ATSC strategies. In addition, ASC/3 SILS concept enables the use of multiple virtual ASC/3 controllers simultaneously. They are compliant with the National Transportation Communications for Intelligent Transportation Systems Protocol (NTCIP) and Transmission Control Protocol / Internet Protocol (TCP/IP). Finally, the ASC/3 controller has built-in TSP features for green extension and red reduction strategies. Custom defined TSP strategies can also be achieved through the logic processor.

### **3.1.2 Traffic Microsimulation**

VISSIM models dynamic and stochastic movements of individual vehicles according to the physical characteristics of different vehicle types, rules of driver behavior, traffic management rules and others. The simulation performance an online animation of the traffic flow and offline reports of traffic performance measurement can be generated. Specifically, it consists of two programs: traffic flow model and

the signal control model. The advantage of splitting simulation into two programs is that the signal control strategy implemented in VISSIM can be flexible. VISSIM provides the module of External Signal Control, which enables to simulate user developed signal control strategies as a separate application (\*.exe) or a program library (\*.dll). The current VISSIM software provides an Econolite ASC/3 module to simulate signalized intersections using ASC/3 controller. Dedicated user interface is available to set control parameters.

### **3.1.3 Date Flow and Integration**

Figure 3.1 shows a real world transportation environment with actuated control. There is a two-way communication. The detector calls are sent to the traffic controller, which processes the inputs through its traffic control logic and returns current signal states to signal heads through the relay. The vehicles react to the traffic light by stopping in case of red signal, or proceeding in case of green signal. In SILS, the traffic signal control logic is implemented with virtual traffic signal controller, and the real traffic environment is replaced by the microsimulation software. The vehicles in the simulation environment generate detector calls, which are sent to the virtual controller. The controllers updated the phase status according the programmed logic. The updated phase status is subsequently sent back to the simulated environment.

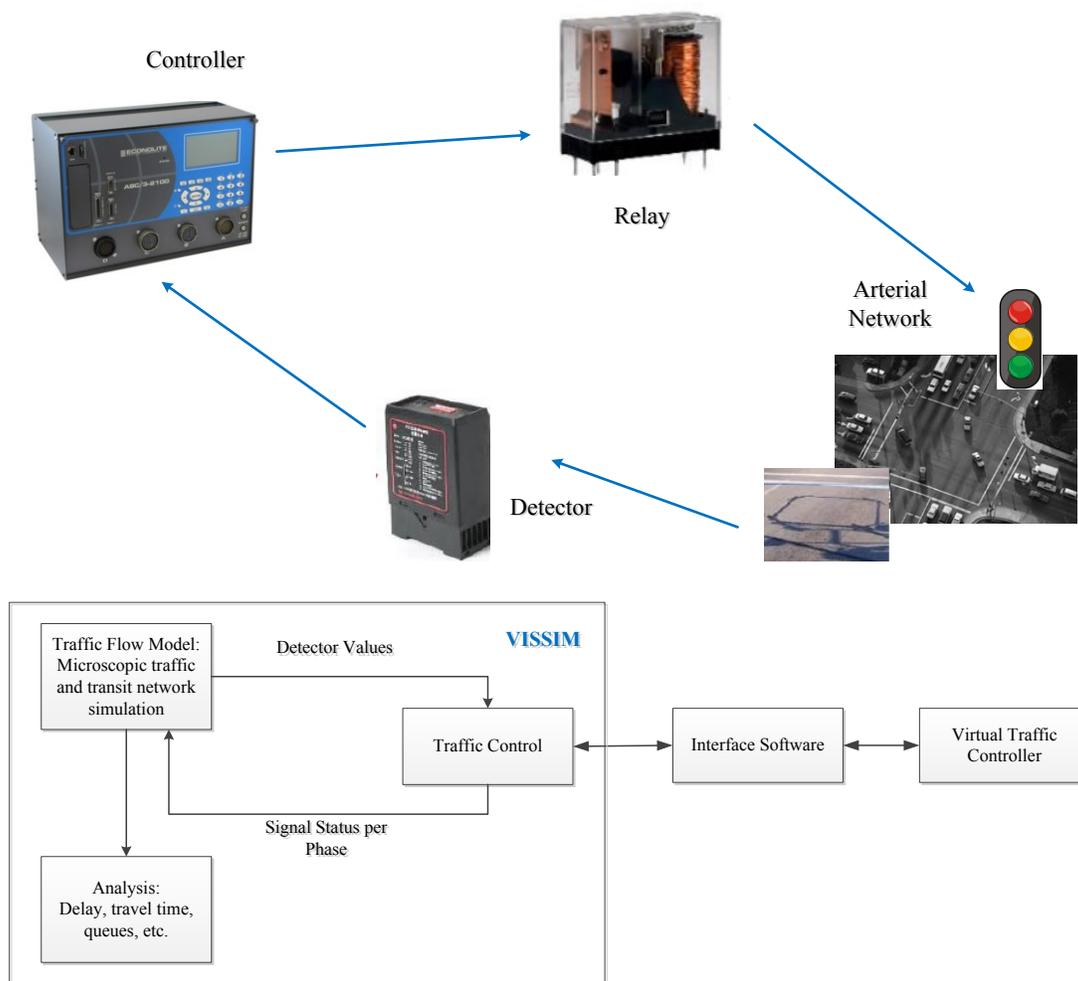


Figure 3.1 Real Transportation Environment and SILS Data Flow

The ASC/3 SILS has several components: the Data Manager, Traffic Control Kernel, Controller Front Panel Simulator, and VISSIM DLL Interface components [85].

- Data Manager manages the timing data in the Windows environment. The database file is identical as an actual ASC/3 controller. Any changes in the controller settings are stored in the database.

- Traffic Control Kernel, acting as the virtual ASC/3 core software, encompasses all internal processing that occurs between the mapped field inputs and commanded field outputs.
- The Controller Front Panel Simulator is a Graphical User Interface (GUI) designed to simulate the keypad of a physical controller.
- The VISSIM DLL allows VISSIM to pass Input/Output functions to the virtual controllers and to receive signal timing data back.

### **3.2 Simulation Platform Architecture**

Although ASC/3 SILS provides the ability to model different signal timing strategies, ATSC strategy requires extending the built in capabilities to implement the optimized decisions using the native controller functions. In addition, the platform must provide efficient and reliable communication amongst adaptive control actions and ASC/3 SILS. Figure 3.2 illustrates how the adaptive control strategies are implemented in the applied simulation platform. It contains ASC/3 SILS, ASC/3 interface, and control system. Signal timing data are imported through NTCIP from ASC/3 interface to the control system. Traffic performance measurement data are imported from VISSIM to the control system. A Microsoft Visual C++ application is created to control the simulation process and continuously read VISSIM evaluation files by using the Component Object Model (COM) interface. The signal timings are then modified through the adaptive algorithms in the control system. Finally, the new optimized signal timings are sent back to the ASC/3 SILS. Optimization of ASC/3

controller signal timings is achieved by the interface between the control system and ASC/3 SILS, and is evaluated through VISSIM simulations.

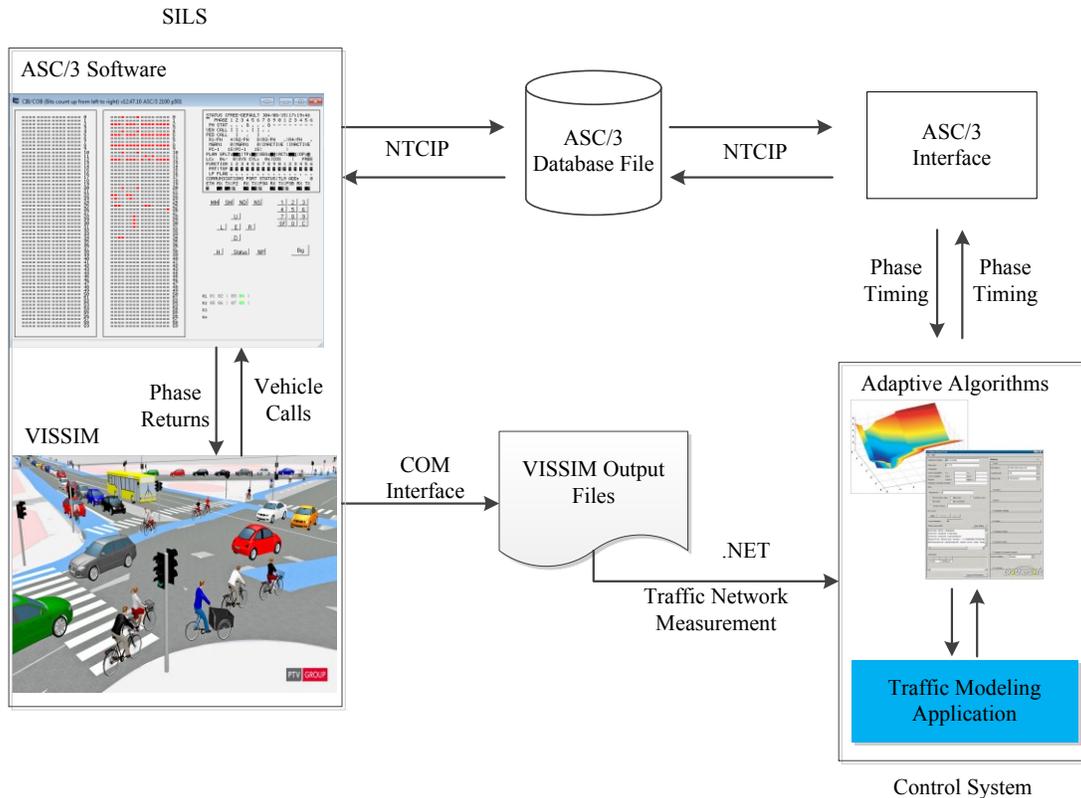


Figure 3.2 Framework of Simulation Platform

### 3.2.1 ASC/3 Interface

ASC/3 interface connects the virtual controller and control system, and it enables to read and override the control logic of the controller by sending the appropriate actions at the appropriate time stamps. The transfers follow NTCIP and communicate through an Ethernet port via Simple Network Management Protocol (SNMP). According to the SNMP protocols, every data frame includes an opening flag, address, control, information, cyclic redundancy check, and a closing flag.

Different frame types are transmitted to different units at every second or 0.1 second. The ASC/3 interface signal timing data through extracting the *address*, *control*, and *information* fields from the controller databases. Then the data are decrypted to specific signal timings recognizable by the control system. Finally, the interface encrypts the optimized signal timing plans back into the ASC/3 database format.

### **3.2.2 Control System**

The control system includes adaptive optimization program, optimization program and traffic flow modeling. To consider the impacts of signal timing plans on traffic flow dynamic, the traffic flow modeling is applied to predict the cycle-based traffic flow states based on loop detector data. The inputs to the optimization models are predicted traffic flow dynamic and signal status. The outputs from the optimization models are adaptive control strategies. Genetic Algorithm (GA) is used to solve the problem. Many previous studies have also shown the effectiveness of GA when solving signal optimization problems [86-88].

### **3.2.3 Data Flow and Integration**

Figure 3.3 describes the interaction of the different components. After the initializations, two major threads are active. One thread is responsible for servicing the commands from the controller. It also updates appropriate variables based on information available in the controller commands. The other thread services the data streams from the simulation software. These streams can include request for traffic data and updates to the loop detector calls.

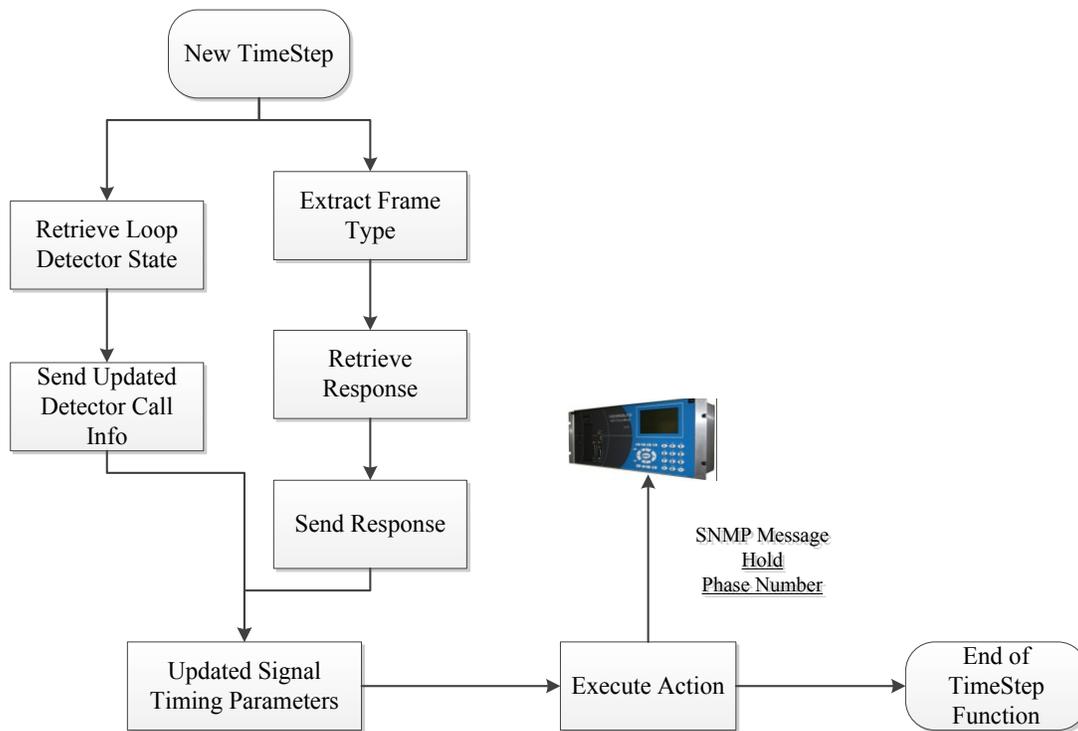
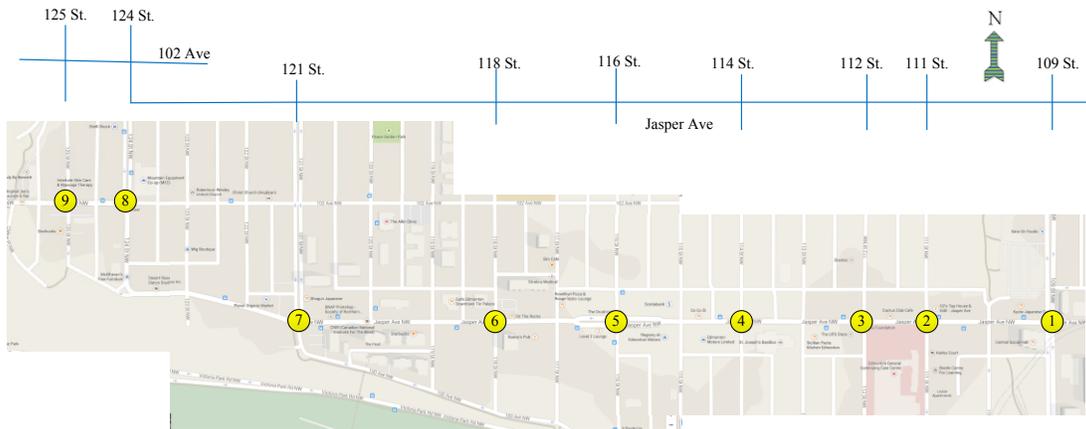


Figure 3.3 Data Flowchart

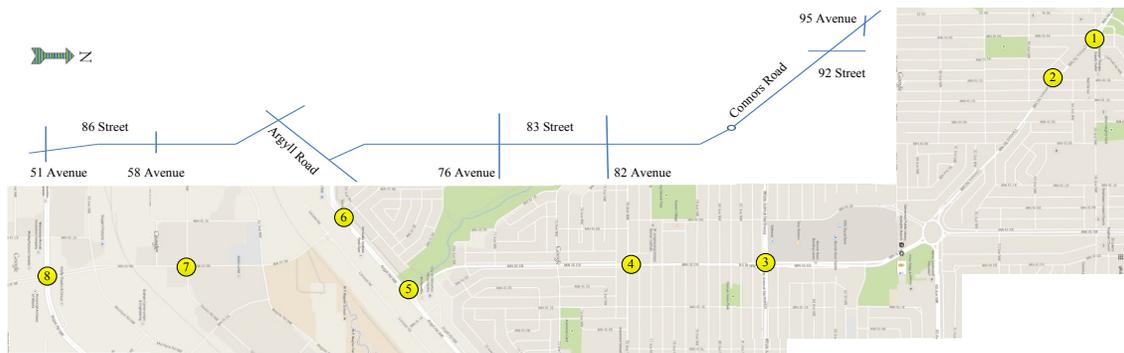
### 3.3 Test-network Simulation Model

#### 3.3.1 Study Corridors

To evaluate the performance of the proposed model, this study has selected two arterial corridors in the city of Edmonton, Alberta, Canada. As shown in Figure 3.4, the downtown corridor is about 2.3 km long and consists of 9 intersections, and the southeast corridor is about 7.4 km long and consists of 8 signalized intersections. The downtown corridor with short link length stretches along the Jasper Ave from 109 Street to 125 Street. The southeast corridor with long link length stretches along the 83rd Street, Argyll Road, the 86th Street, and the Connors Road. The PM peak period was selected for simulation because it has the highest volume of ridership.



(a) Downtown Corridor



(b) Southeast Corridor

Figure 3.4 Study Corridors

The signal timing plan is shown in Table 3.1 and 3.2. The signal plans are represented following standard NEMA 8-phases diagram. For southeast corridor, Intersection (Int.) 4, Int. 7 and Int. 8 have a standard four-phase plan without protected left-turn phases. The remaining intersections contain protected left-turn phases. For the downtown corridor, Int. 3, Int. 4, Int. 6, Int. 7, and Int. 9 have a standard four-phase plan without protected left-turn phases. The remaining intersections contain protected left-turn phases. Coordinated Phase 2 and Phase 6 are

the reference phases. Yellow time and all-red time were set according to the respective real signal timing.

Table 3.1 Signal Timings at Downtown Corridor

Intersection No.	Cycle	Offset	Timing Plan								
1	110	70	<table border="1"> <tr> <td><math>\phi^1</math> 6s</td> <td><math>\phi^2</math> 42s</td> <td><math>\phi^3</math> 9s</td> <td><math>\phi^4</math> 37s</td> </tr> <tr> <td><math>\phi^5</math> 6s</td> <td><math>\phi^6</math> 42s</td> <td colspan="2"><math>\phi^8</math> 46s</td> </tr> </table>	$\phi^1$ 6s	$\phi^2$ 42s	$\phi^3$ 9s	$\phi^4$ 37s	$\phi^5$ 6s	$\phi^6$ 42s	$\phi^8$ 46s	
$\phi^1$ 6s	$\phi^2$ 42s	$\phi^3$ 9s	$\phi^4$ 37s								
$\phi^5$ 6s	$\phi^6$ 42s	$\phi^8$ 46s									
2	110	9	<table border="1"> <tr> <td colspan="2"><math>\phi^2</math> 67s</td> <td rowspan="2"><math>\phi^4</math> 30s</td> </tr> <tr> <td><math>\phi^5</math> 6s</td> <td><math>\phi^6</math> 61s</td> </tr> </table>	$\phi^2$ 67s		$\phi^4$ 30s	$\phi^5$ 6s	$\phi^6$ 61s			
$\phi^2$ 67s		$\phi^4$ 30s									
$\phi^5$ 6s	$\phi^6$ 61s										
3	110	10	<table border="1"> <tr> <td><math>\phi^2</math> 68s</td> <td><math>\phi^4</math> 32s</td> </tr> <tr> <td><math>\phi^6</math> 68s</td> <td><math>\phi^8</math> 32s</td> </tr> </table>	$\phi^2$ 68s	$\phi^4$ 32s	$\phi^6$ 68s	$\phi^8$ 32s				
$\phi^2$ 68s	$\phi^4$ 32s										
$\phi^6$ 68s	$\phi^8$ 32s										
4	110	28	<table border="1"> <tr> <td><math>\phi^2</math> 74s</td> <td><math>\phi^4</math> 26s</td> </tr> <tr> <td><math>\phi^6</math> 74s</td> <td><math>\phi^8</math> 26s</td> </tr> </table>	$\phi^2$ 74s	$\phi^4$ 26s	$\phi^6$ 74s	$\phi^8$ 26s				
$\phi^2$ 74s	$\phi^4$ 26s										
$\phi^6$ 74s	$\phi^8$ 26s										
5	110	60	<table border="1"> <tr> <td><math>\phi^2</math> 49s</td> <td><math>\phi^3</math> 7s</td> <td><math>\phi^4</math> 38s</td> </tr> <tr> <td><math>\phi^5</math> 7s</td> <td><math>\phi^6</math> 42s</td> <td><math>\phi^7</math> 7s</td> <td><math>\phi^8</math> 38s</td> </tr> </table>	$\phi^2$ 49s	$\phi^3$ 7s	$\phi^4$ 38s	$\phi^5$ 7s	$\phi^6$ 42s	$\phi^7$ 7s	$\phi^8$ 38s	
$\phi^2$ 49s	$\phi^3$ 7s	$\phi^4$ 38s									
$\phi^5$ 7s	$\phi^6$ 42s	$\phi^7$ 7s	$\phi^8$ 38s								
6	110	50	<table border="1"> <tr> <td><math>\phi^2</math> 69s</td> <td><math>\phi^4</math> 31s</td> </tr> <tr> <td><math>\phi^6</math> 69s</td> <td><math>\phi^8</math> 31s</td> </tr> </table>	$\phi^2$ 69s	$\phi^4$ 31s	$\phi^6$ 69s	$\phi^8$ 31s				
$\phi^2$ 69s	$\phi^4$ 31s										
$\phi^6$ 69s	$\phi^8$ 31s										
7	110	84	<table border="1"> <tr> <td><math>\phi^2</math> 71s</td> <td rowspan="2">Fixed Ped Jump</td> <td><math>\phi^4</math> 26s</td> </tr> <tr> <td><math>\phi^6</math> 71s</td> <td><math>\phi^8</math> 26s</td> </tr> </table>	$\phi^2$ 71s	Fixed Ped Jump	$\phi^4$ 26s	$\phi^6$ 71s	$\phi^8$ 26s			
$\phi^2$ 71s	Fixed Ped Jump	$\phi^4$ 26s									
$\phi^6$ 71s		$\phi^8$ 26s									
8	110	89	<table border="1"> <tr> <td><math>\phi^1</math> 5s</td> <td><math>\phi^2</math> 24s</td> <td><math>\phi^3</math> 44s</td> <td><math>\phi^4</math> 27s</td> </tr> <tr> <td colspan="2"><math>\phi^6</math> 29s</td> <td colspan="2"><math>\phi^8</math> 71s</td> </tr> </table>	$\phi^1$ 5s	$\phi^2$ 24s	$\phi^3$ 44s	$\phi^4$ 27s	$\phi^6$ 29s		$\phi^8$ 71s	
$\phi^1$ 5s	$\phi^2$ 24s	$\phi^3$ 44s	$\phi^4$ 27s								
$\phi^6$ 29s		$\phi^8$ 71s									
9	110	5	<table border="1"> <tr> <td><math>\phi^2</math> 84s</td> <td><math>\phi^4</math> 16s</td> </tr> <tr> <td><math>\phi^6</math> 84s</td> <td><math>\phi^8</math> 16s</td> </tr> </table>	$\phi^2$ 84s	$\phi^4$ 16s	$\phi^6$ 84s	$\phi^8$ 16s				
$\phi^2$ 84s	$\phi^4$ 16s										
$\phi^6$ 84s	$\phi^8$ 16s										

Table 3.2 Signal Timings at Southeast Corridor

Intersection No	Cycle	Offset	Timing Plan								
1	100	79	<table border="1"> <tr> <td><math>\phi_2</math> 90s</td> <td><math>\phi_4</math> 10s</td> </tr> <tr> <td><math>\phi_5</math> 29s</td> <td><math>\phi_6</math> 61s</td> </tr> <tr> <td></td> <td><math>\phi_8</math> 10s</td> </tr> </table>	$\phi_2$ 90s	$\phi_4$ 10s	$\phi_5$ 29s	$\phi_6$ 61s		$\phi_8$ 10s		
$\phi_2$ 90s	$\phi_4$ 10s										
$\phi_5$ 29s	$\phi_6$ 61s										
	$\phi_8$ 10s										
2	100	0	<table border="1"> <tr> <td><math>\phi_2</math> 12s</td> <td><math>\phi_4</math> 88s</td> </tr> <tr> <td><math>\phi_6</math> 12s</td> <td><math>\phi_8</math> 88s</td> </tr> </table>	$\phi_2$ 12s	$\phi_4$ 88s	$\phi_6$ 12s	$\phi_8$ 88s				
$\phi_2$ 12s	$\phi_4$ 88s										
$\phi_6$ 12s	$\phi_8$ 88s										
3	100	96	<table border="1"> <tr> <td><math>\phi_2</math> 69s</td> <td><math>\phi_4</math> 31s</td> </tr> <tr> <td><math>\phi_5</math> 17s</td> <td><math>\phi_6</math> 52s</td> </tr> <tr> <td></td> <td><math>\phi_8</math> 31s</td> </tr> </table>	$\phi_2$ 69s	$\phi_4$ 31s	$\phi_5$ 17s	$\phi_6$ 52s		$\phi_8$ 31s		
$\phi_2$ 69s	$\phi_4$ 31s										
$\phi_5$ 17s	$\phi_6$ 52s										
	$\phi_8$ 31s										
4	50	44	<table border="1"> <tr> <td><math>\phi_2</math> 34s</td> <td><math>\phi_4</math> 16s</td> </tr> <tr> <td><math>\phi_6</math> 34s</td> <td><math>\phi_8</math> 16s</td> </tr> </table>	$\phi_2$ 34s	$\phi_4$ 16s	$\phi_6$ 34s	$\phi_8$ 16s				
$\phi_2$ 34s	$\phi_4$ 16s										
$\phi_6$ 34s	$\phi_8$ 16s										
5	100	24	<table border="1"> <tr> <td><math>\phi_2</math> 23s</td> <td><math>\phi_3</math> 54s</td> <td><math>\phi_4</math> 23s</td> </tr> <tr> <td></td> <td><math>\phi_8</math> 77s</td> <td></td> </tr> </table>	$\phi_2$ 23s	$\phi_3$ 54s	$\phi_4$ 23s		$\phi_8$ 77s			
$\phi_2$ 23s	$\phi_3$ 54s	$\phi_4$ 23s									
	$\phi_8$ 77s										
6	100	92	<table border="1"> <tr> <td><math>\phi_2</math> 28s</td> <td><math>\phi_3</math> 8s</td> <td><math>\phi_4</math> 64s</td> </tr> <tr> <td><math>\phi_6</math> 28s</td> <td><math>\phi_7</math> 33s</td> <td><math>\phi_8</math> 39s</td> </tr> </table>	$\phi_2$ 28s	$\phi_3$ 8s	$\phi_4$ 64s	$\phi_6$ 28s	$\phi_7$ 33s	$\phi_8$ 39s		
$\phi_2$ 28s	$\phi_3$ 8s	$\phi_4$ 64s									
$\phi_6$ 28s	$\phi_7$ 33s	$\phi_8$ 39s									
7	50	8	<table border="1"> <tr> <td><math>\phi_2</math> 29s</td> <td><math>\phi_4</math> 21s</td> </tr> <tr> <td><math>\phi_6</math> 29s</td> <td><math>\phi_8</math> 21s</td> </tr> </table>	$\phi_2$ 29s	$\phi_4$ 21s	$\phi_6$ 29s	$\phi_8$ 21s				
$\phi_2$ 29s	$\phi_4$ 21s										
$\phi_6$ 29s	$\phi_8$ 21s										
8	100	90	<table border="1"> <tr> <td><math>\phi_1</math> 10s</td> <td><math>\phi_2</math> 33s</td> <td><math>\phi_3</math> 18s</td> <td><math>\phi_4</math> 39s</td> </tr> <tr> <td><math>\phi_5</math> 17s</td> <td><math>\phi_6</math> 26s</td> <td><math>\phi_7</math> 14s</td> <td><math>\phi_8</math> 43s</td> </tr> </table>	$\phi_1$ 10s	$\phi_2$ 33s	$\phi_3$ 18s	$\phi_4$ 39s	$\phi_5$ 17s	$\phi_6$ 26s	$\phi_7$ 14s	$\phi_8$ 43s
$\phi_1$ 10s	$\phi_2$ 33s	$\phi_3$ 18s	$\phi_4$ 39s								
$\phi_5$ 17s	$\phi_6$ 26s	$\phi_7$ 14s	$\phi_8$ 43s								

### 3.3.2 Modeling Process

The study corridors were modeled in VISSIM simulation model with existing network geometry, traffic volumes, turning movements at intersections, signal timing data, and transit operations data. The VISSIM model of two corridors was carefully calibrated and validated to resemble field conditions as much as possible. The modeling process started from the basic network geometry. After that, the geometry was fine-tuned, all traffic and transit data incorporated, and the model was calibrated and validated. The data coded in the model were based on real data collected in the

field or from City of Edmonton. The flowchart of the modeling process is given in Figure 3.5.

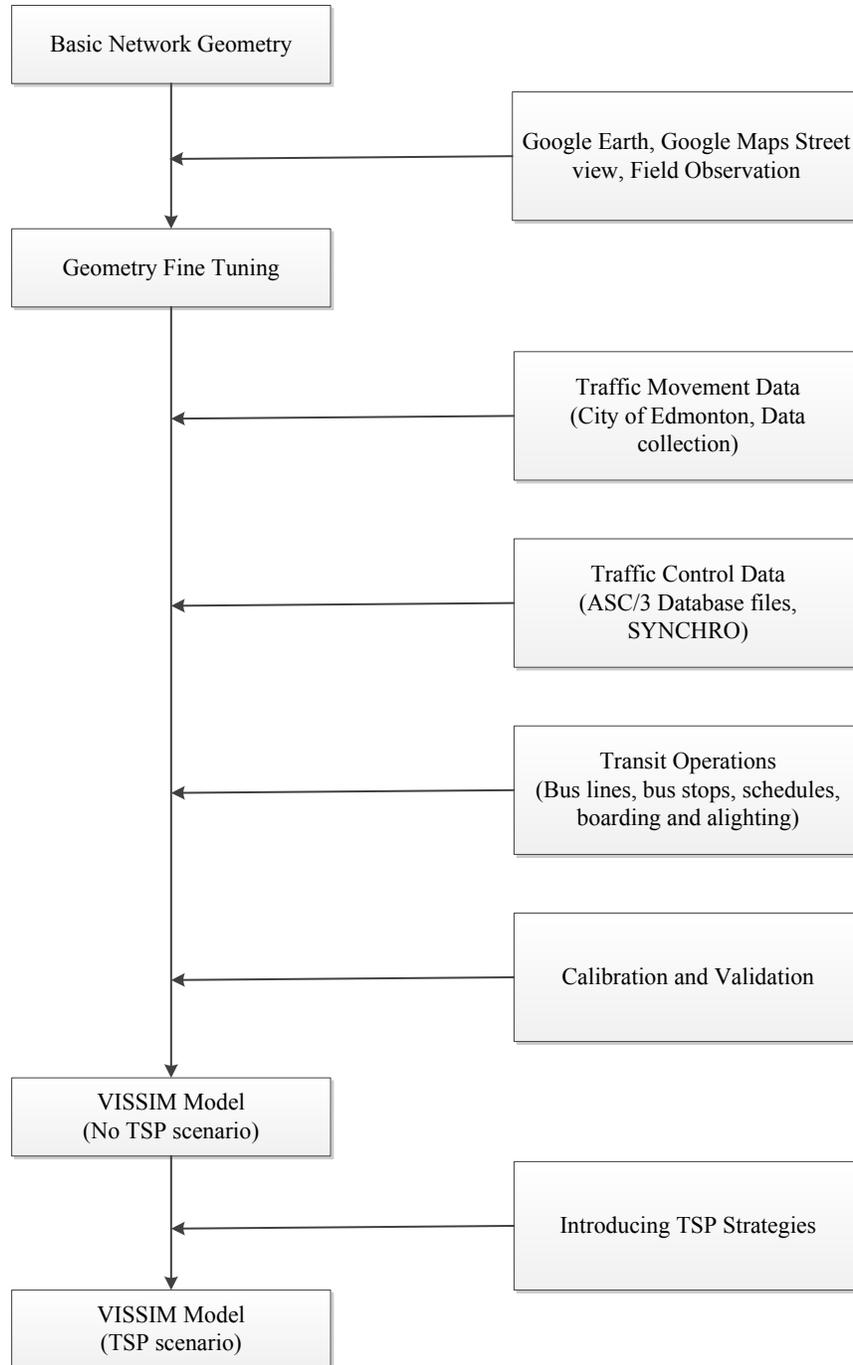
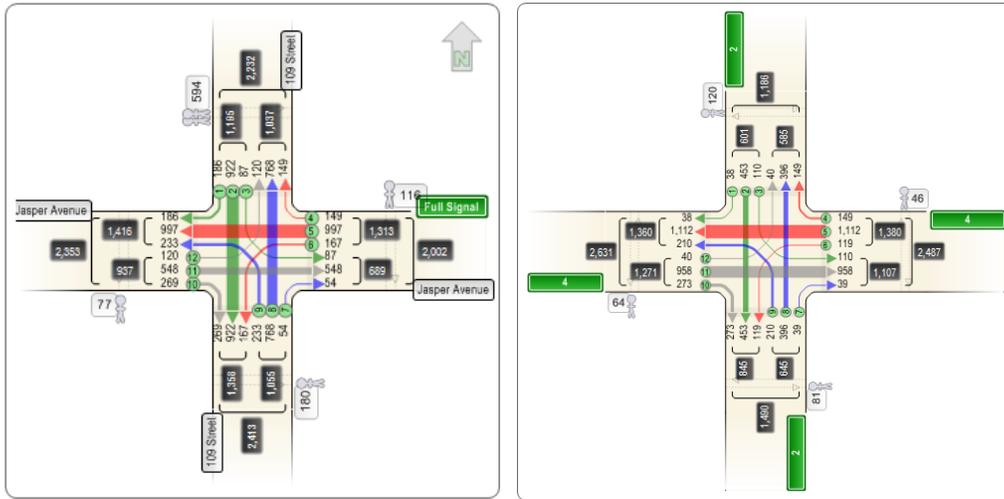


Figure 3.5 Modeling Process Flowchart

The intersection traffic counts were taken at 5-minute intervals, as shown in Table 3.3. Although the turning movement counts were not collected on same date, they represent the real-world traffic pattern. It is necessary to convert the data to the same time period by traffic balancing. The raw data contains not only turning movement counts, but also vehicle types. Therefore, the average heavy vehicle rate can be calculated and modeled in VISSIM. Pedestrian counts are also included in the raw data. Figure 3.6 gives examples of the hourly turning movement and the lane assignment at each intersection.

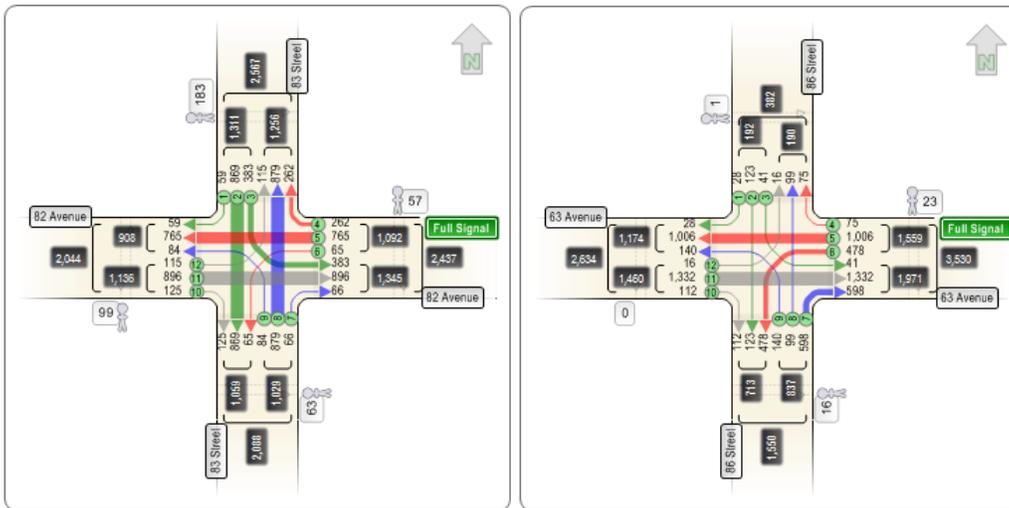
Table 3.3 Turning Movements Data

Downtown Int. No.	Data Missing	Interval	Collection Date	Southeast Int. No.	Data Missing	Interval	Collection Date
1	No	5 min	11/08/2011	1	No	5 min	05/30/2011
2	No	5 min	04/30/2009	2	No	5 min	05/27/2009
3	No	5 min	05/04/2009	3	No	5 min	04/19/2011
4	No	5 min	05/07/2009	4	No	5 min	02/15/2011
5	No	5 min	05/06/2009	5	No	5 min	09/16/2010
6	No	5 min	05/25/2009	6	No	5 min	09/16/2010
7	No	5 min	05/26/2009	7	No	5 min	09/15/2010
8	No	5 min	05/25/2009	8	No	5 min	09/15/2010
9	No	5 min	05/26/2009				



(a) Int. 1 of downtown corridor

(b) Int. 5 of downtown corridor



(c) Int. 4 of southeast corridor

(d) Int. 7 of southeast corridor

Figure 3.6 Examples of Turning Movement at Intersections

Based on bus stop detail document, bus stops location and bus stop length are paced on the VISSIM simulation road networks. Then, the bus schedule of each bus line was configured according to the bus departure times and headways described in the September bus schedule provided by ETS. The average headways of major bus routes and minor bus routes are 10 minutes and 15 minutes, respectively, during peak

hours. VISSIM allows users to set up both side-street transit stations and bus bays; the bus stop type was determined using Google Maps. For downtown corridor, four lines travel the whole corridor: No. 1, No. 5, No. 120, and No. 135. For the southeast corridor, nine lines travel the whole transit corridor. Those lines are No. 8, No. 15, No. 61, No. 64, No. 65, No. 66, No. 68, No. 69, No. 72. Only No. 8 and No. 15 operate all the time, while the other routes are express only for peak hours. The bus dwell time is important when estimating bus link travel times. This study uses empirical bus dwell time calculation according to the number of boarding passengers, alighting times, and bus clearance times at bus stops. Ridership estimations for all lines were provided by ETS and they were coded in the model.

The signal timings and the detector locations were built in VISSIM according to documents provided by the City of Edmonton's traffic operations branch. Three key configurations are needed to realize the TSP function in VISSIM, including configuring the detectors in VISSIM to detect TSP requests, developing the TSP plans in VISSIM and mapping the bus detectors in VISSIM. To distinguish the TSP-enabled buses from the general traffic and regular buses, a new vehicle class is defined in simulation as 'TBus'. The check-in and check-out detectors would send a pulsed signal to the signal controller only when a TSP-enabled bus passes the detectors. Once a pulsed signal is received by the ASC/3 controller, the TSP signal timing will override the existing timing without interrupting the coordination. Two major parameters need to be configured first: MAX RDTN which is the maximum time that other phases can be reduced during priority and MAX EXTN which is the maximum time a phase can be extended during priority.

### **3.3.3 Calibration and Validation**

The calibration and validation are important for the simulation evaluation considering the following reasons. First, testing the new traffic signal control strategies prior to field implementation is essential under representative traffic conditions. Confounding effects hinder before-and-after field tests. Controlled simulation experiments can draw strong statistical conclusions. Second, simulation saves time, effort and costs induced by testing on a field controller. Third, after the simulation test, the adaptive control strategies can be easily transferred to the field controllers.

Traffic movements for each signalized intersection were used to calibrate traffic operations in the model. The field data were collected in different time periods and different days, which impacted the precision of the counts. These traffic flows needed to be adjusted to account for the unbalanced traffic counts. For this purpose, some additional traffic generators and collectors were used for inflow or outflow of the additional traffic. Calibration was performed by iteratively adjusting traffic counts in the model until a highly correlated match between the field data and the modeled data was reached. The coefficient of determination,  $R^2$ , analysis was conducted to investigate the relationship between simulation output and field observation.  $R^2$  indicates how well data points fit a proposed line or curve. On the graph, the actual traffic volumes are represented on the y-axis and simulated traffic volumes are represented on the x-axis: if the actual traffic volumes exactly match the simulated volumes, then all data points would be on a 45 degree line and  $R^2$  value for

the dataset would equal 1. However, the farther away from the 45 degree line those data points are, the greater the discrepancies between actual and simulated traffic volumes, and hence, the lower the  $R^2$  value of the dataset. The expression for  $R^2$  is:

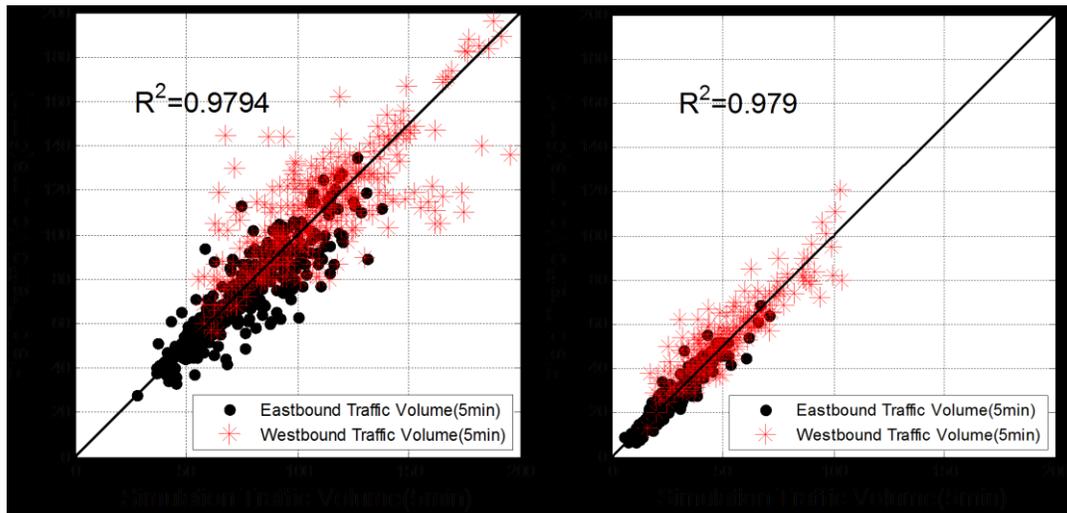
$$R^2 = 1 - \frac{\sum_i (v_i - f_i)^2}{\sum_i (v_i - \frac{1}{n} \sum_{i=1}^n v_i)^2} \quad (3-1)$$

Where:

$v_i$ : the 5-minute traffic volumes from the VISSIM simulation

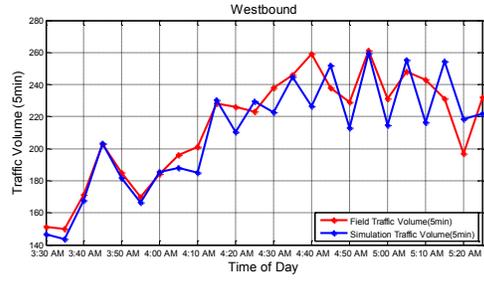
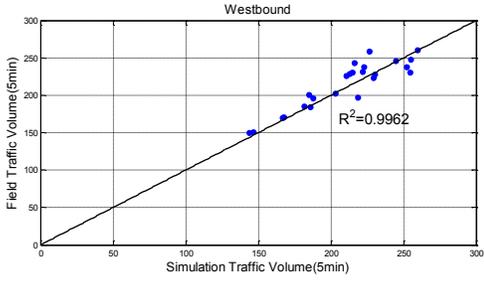
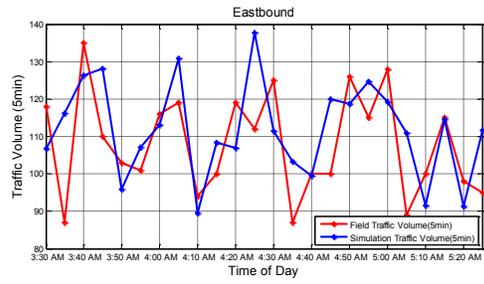
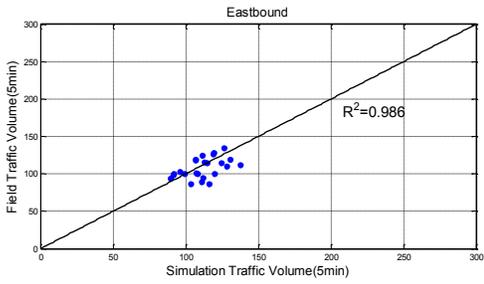
$f_i$ : the 5-minute empirical traffic volumes

In addition, the simulation output is a mean value through 10 times of run. Thus, the variation caused by random factors can be eliminated. Figure 3.7 shows results from the calibration process. High  $R^2$  values indicate a high correlation between the data sets collected in the field and those from the simulation.

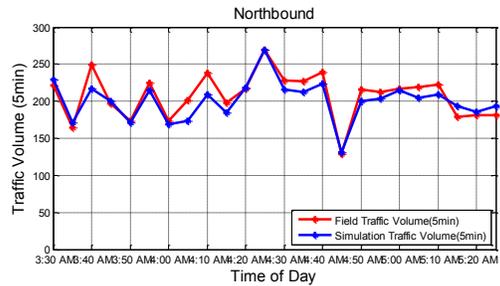
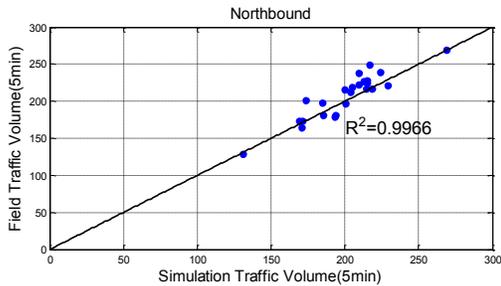
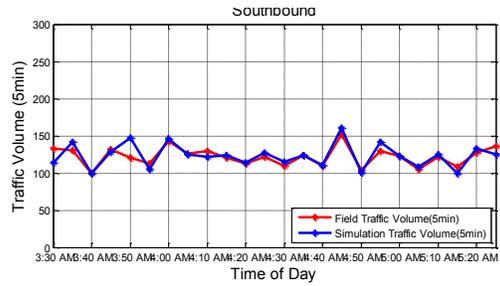
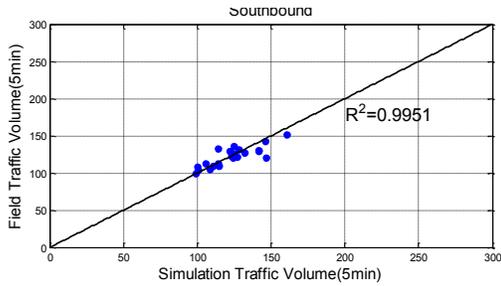


(a) Downtown Corridor

(b) Southeast Corridor



(c) Downtown Segment



(d) Southeast Segment

Figure 3.7 Model Calibration Results

To validate the model, bus travel times from the field were compared with those from the model. The process was performed iteratively by setting speed limits,

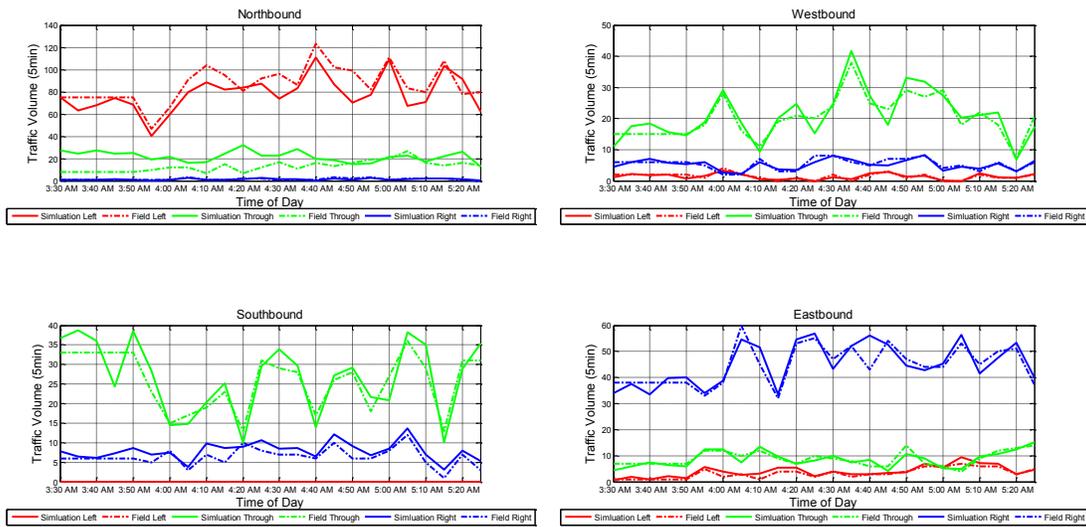
speed distributions, and driving behavior in VISSIM. Table 3.4 shows a comparison of the two sets of bus travel times, averaged from ten simulation runs over a 2-hour peak period.

Table 3.4 Model Validation Results

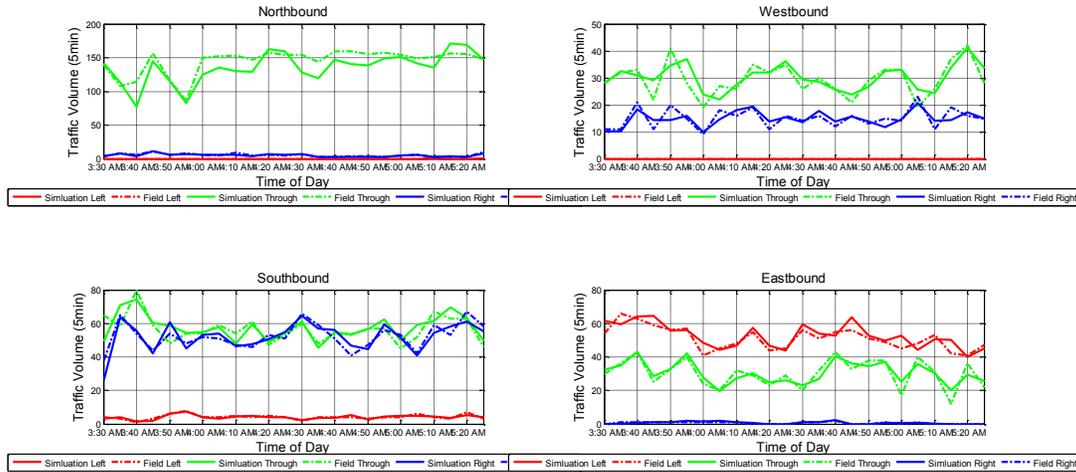
	Downtown Corridor		Southeast Corridor	
	EB (s)	WB (s)	SB (s)	NB (s)
Field Observation	451.8	559.2	1105.0	1178.0
Simulation	437.4	538.8	1089.7	1068.6
Relative Error	-3.1%	-3.6%	-1.38%	-9.28%

### 3.3.4 Design of Experiments

The PM peak period, from 15:30 to 17:30, was selected for simulation. It is necessary to input some vehicles before evaluation. This is the warm-up time, which is 10 minutes. Cool down time is also necessary, which is also 10 minutes. During the simulation, the rates at which vehicles are assumed to enter the control area at the boundary of the control area vary over time according to Figure 3.8.



### (a) Downtown Corridor



### (b) Southeast Corridor

Figure 3.8 Temporal Variations of Traffic Demand

This study did not attempt to answer the question of how this system was compared to other adaptive systems in the market. Instead, the analysis was conducted to evaluate the performance of proposed models with the actuated control. The base case model involved the existing traffic conditions for the PM peak period. VISSIM models were developed, calibrated and validated for current traffic conditions. Here we applied another scenario, with some small changes in traffic demands to make them more suitable for the focus of the research. By using the actual PM peak volume as the base line, this study generated two possible levels of traffic demand conditions: (1) current PM peak volume; (2) 15% increase of current volume. These corresponded to different levels of intersection saturation, which were estimated by the intersection capacity utilization in SYNCHRO. Table 3.5 showed the saturation rate of all intersections under the different levels of traffic demand.

Table 3.5 Intersection Saturation Rate under Different Traffic Demand Conditions

Downtown Corridor	Saturation Rate		Southeast Corridor	Saturation Rate	
	Current	15%		Current	15%
Intersection 1	0.71	0.82	Intersection 1	0.79	0.90
Intersection 2	0.80	0.93	Intersection 2	0.83	0.98
Intersection 3	0.81	0.93	Intersection 3	1.08	1.23
Intersection 4	0.75	0.89	Intersection 4	1.01	1.23
Intersection 5	0.89	1.12	Intersection 5	0.63	0.85
Intersection 6	0.66	0.80	Intersection 6	0.81	0.96
Intersection 7	0.83	0.96	Intersection 7	0.75	0.89
Intersection 8	0.73	0.92	Intersection 8	0.81	0.93
Intersection 9	0.90	1.06			

# **Chapter 4 Proactive Arterial Signal Optimization with Embedded Enhanced Store-and-Forward Model**

## **4.1 Introduction**

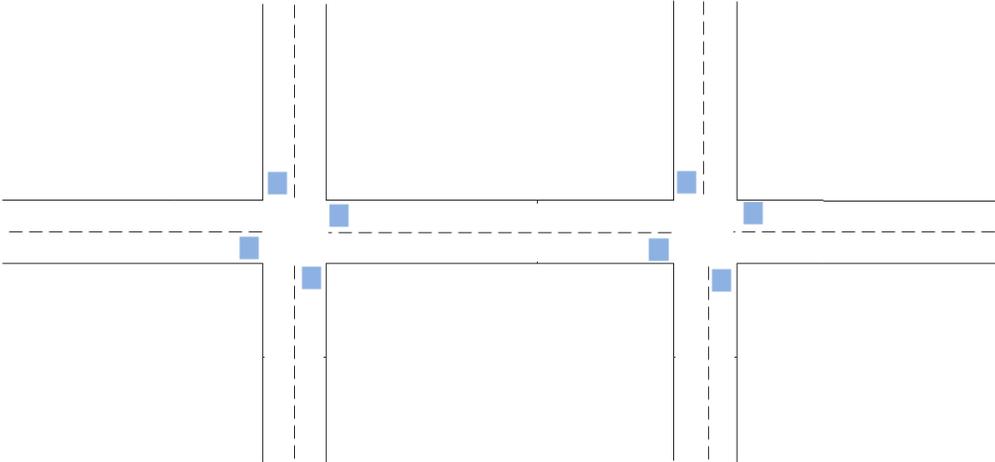
A number of elaborate traffic flow models, which are deductively derived to describe the complex interactions between traffic states evolution and key control parameters, have been applied to provide relatively accurate predictions [12-14, 39, 64, 89-91]. Subsequently, a number of model-based proactive control systems have been presented in literature or implemented in the field, some of which are proven effective in practice [16-18, 34, 60, 61]. However, it remains a challenging task to generate applicable and reliable network-wide proactive control system using traffic flow models: efficiency must be improved, and it is important to find a balance between accuracy and complexity. Furthermore, another challenging task is to generate reliable signal timing plans that can systematically and globally consider dynamic queue interactions among different lanes and adjacent intersections in congested traffic conditions.

A particular simplified control design pursued by various works in the past is based on the store-and-forward model (SFM) [39, 64]. SFM enables the mathematical description of the traffic flow process without use of discrete variables, allowing for efficient optimization and control methods in real-time for a large-scale network [16]. SFM has several obvious limitations, which have not been thoroughly

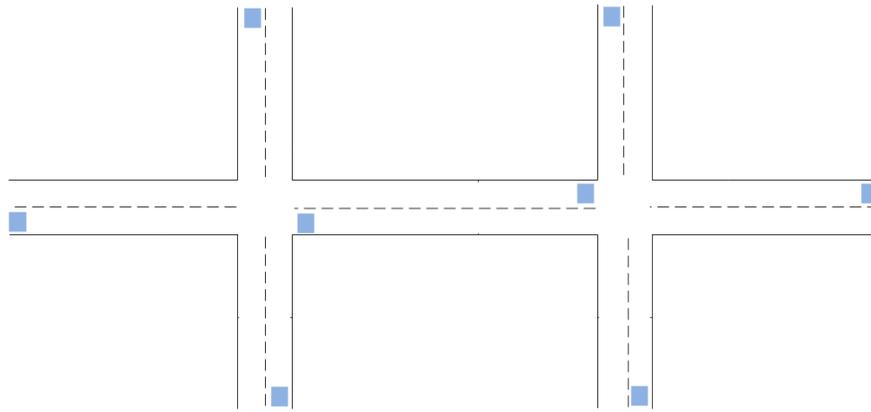
studied. First, SFM considers the constant travel time in one specific link and cannot reveal the queue dynamic in one cycle. Second, SFM assumes that all movements (straight and right- and left-turning) of an incoming link receive the right of way simultaneously, so SFM has difficulty integrating with multiple signal phases. Third, it is very common to see queue interaction among neighboring lane groups in a link. Turning vehicles strongly influence the validity of the model. In this study, an enhanced SFM-based signal optimization model is presented to address the aforementioned issues.

**4.2 Enhanced Store-and-Forward Model**

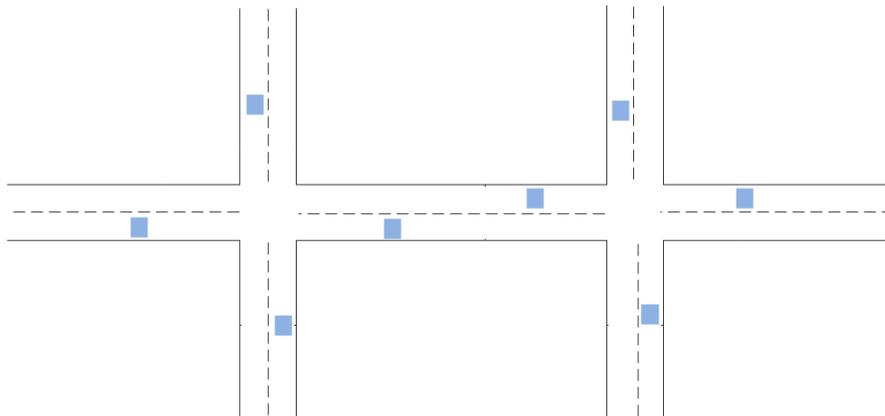
Loop detectors are installed at intersections to collect the required traffic information as the input for control strategies. Figure 4.1 illustrates the loop detector location of several ATSC systems. For instance, SCATS requires loops at stop-line. This study requires loops to be installed at the upstream of the signalized link, similar to the SCOOT configuration.



(a) SCATS



(b) SCOOT



(c) ATCS

Figure 4.1 Detector Requirement for Adaptive Systems

The nature of traffic on the urban network is stop-and-go running condition where vehicles queue at the stop-line during signal red phase whereas, upstream of the link can be free flow. Thus, the speed obtained from the loop cannot be generalized over the signalized link. Similarly, unlike the motorway traffic the occupancy of the loops cannot easily provide the density of the entire link. This section describes the mathematical equations to represent dynamic traffic states for the arterial traffic network. The equations have key features: 1) model traffic flow

evolution along arterial links and nodes; 2) model the merging and diverging of vehicle movements at intersections; 3) capture the physical queue formation and dissipation process; 4) represent the interaction between control parameters and dynamic traffic states. The traffic dynamic includes a process: upstream arrivals, propagation to the end of queue, merging into lane groups, and departing, as shown in Figure 4.2. In order to describe the model, we define  $J$  as the set of nodes (intersections) and  $L$  as the set of links (streets) in the urban traffic network. Link  $j_w$  is marked by its downstream node  $j$  and the direction of west. The sets of links of input flow and output flows for link  $j_w$  are defined as  $I_{j_w}$  and  $O_{j_w}$ .

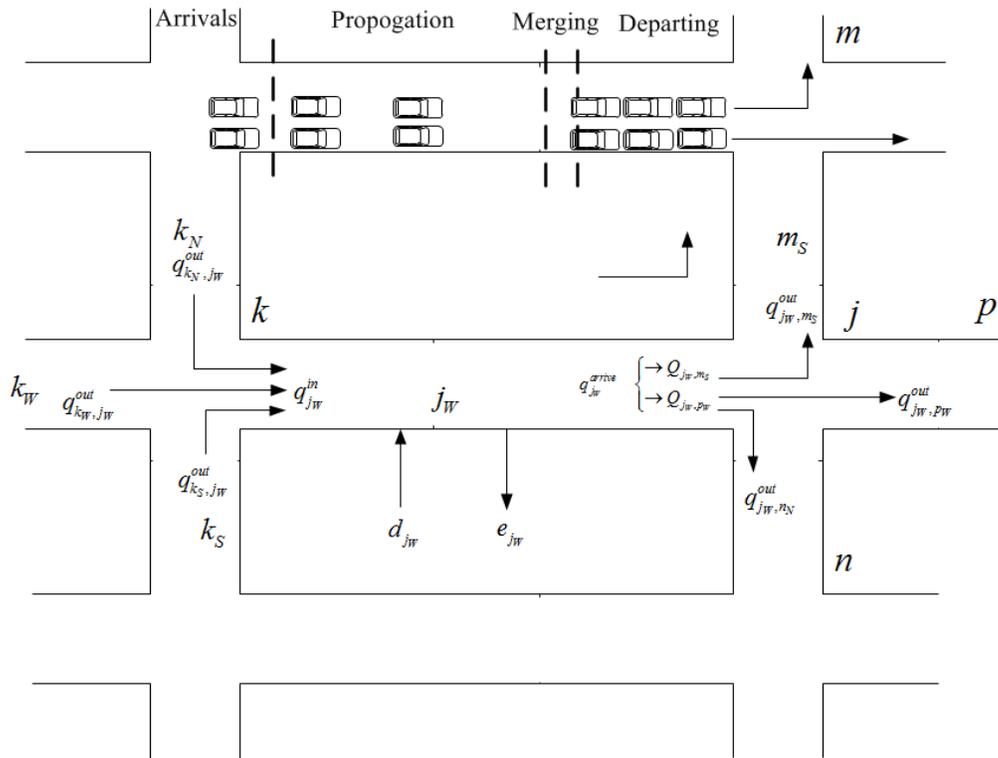


Figure 4.2 Dynamic Traffic Flow Evolutions along Arterial Streets

### A. Upstream Arrivals

Upstream arrival equation describes the flow evolution, which arrives at the upstream of one link over time. Similar to most other research, SFM formulates the inflow to the link  $j_W$  as the sum of departure flows from  $I_{j_W}$ , as shown in Equation (4-1).

$$q_{j_W}^{in}(k) = \sum_{l \in I_{j_W}} q_{l,j_W}^{out}(k) \quad (4-1)$$

Where  $q_{j_W}^{in}(k)$  = the upstream arrival flow of link  $j_W$  during time step  $k$

$q_{l,j_W}^{out}(k)$  = the departing flows from link  $l$  that merge into  $j_W$ , and  $l$  belongs to  $I_{j_W}$

### B. Propagation to the End of Queue

Then upstream arrivals flow propagates to the end of queue. In the SFM, the discrete-time step  $T$  is equal to cycle length. Vehicles entering a link are either stored at the end of this link (during a red signal), or further forwarded to downstream links at the saturation flow rate (during a green signal). Therefore, SFM does not consider the propagation process. Existing other research uses different mathematical equations to model this process.

The Robertson platoon dispersion model takes the form of Equation (4-2) to simulate the the propagation process [49]. However, one critical problem is that if we consider the queue length, then the average link travel time and minimum travel time will be variable in real time.

$$q_{j_W}^{arrive}(t) = F_n * q_{j_W}^{in}(t - P) + (1 - F_n) * q_{j_W}^{arrive}(t - T) \quad (4-2)$$

Where  $q_{j_w}^{arrive}(t)$  = flows arrive at the end of the time  $t$

$t_a$  = the average link travel time

$P$  = the minimum travel time on the link (measured in terms of unit steps

$P = \beta t_a$ )

$q_{j_w}^{in}(t - P)$  = the arrival flow rate at the upstream of the link at time  $t - P$

$F_n$  = the smoothing factor

Liu and Chang's research represents the evolution of upstream arrivals to the end of queue with the average approaching speed [77]. The average speed is depending on the density of the segment between the link upstream and the end of queue, as described by Equation (4-3). Then the number of vehicles arriving at the end of queue is dynamically updated by Equation (4-4).

$$v_{j_w}(k) = \begin{cases} v_{j_w}^{free} & \text{if } \rho_{j_w}(k) < \rho^{\min} \\ v^{\min} + (v_{j_w}^{free} - v^{\min}) \cdot [1 - (\frac{\rho_{j_w}(k) - \rho^{\min}}{\rho^{jam} - \rho^{\min}})^{\alpha}]^{\beta} & \text{if } \rho_{j_w}(k) \in [\rho^{\min}, \rho^{jam}] \\ v^{\min} & \text{if } \rho_{j_w}(k) > \rho^{\min} \end{cases} \quad (4-3)$$

$$q_{j_w}^{arrive}(k) = \min\{\rho_{j_w}(k) \cdot v_{j_w}(k) \cdot N_{j_w}, C_{j_w}(k) - Q_{j_w}(k)\} \quad (4-4)$$

where  $v_{j_w}(k)$  = the average approaching speed;

$\rho^{\min}$  = the minimum critical density below which traffic moves at free flow

speed  $v_{j_w}^{free}$ ;  $v^{\min}$  = the minimum traffic flow speed corresponding to the jam

density ( $\rho^{jam}$ );

$q_{j_w}^{arrive}(k)$  = flows arrive at the tail of the queue during time step  $k$

$C_{j_w}$  = capacity of link  $j_w$ , number of vehicles;

$Q_{j_w}(k)$  = queue length, number of vehicles;

$N_{j_w}$  = number of lanes

$\alpha, \beta$  = constant model parameters to be calibrated.

This study also uses the concept of average approaching speed to represent the propagation process, but the arriving flow at the end of queue at link  $j_w$  is stated as:

$$q_{j_w}^{arrive}(k) = (1 - \alpha(k)) \cdot q_{j_w}^{in}(k - \beta(k)) + \alpha(k) \cdot q_{j_w}^{in}(k - \beta(k) - 1) \quad (4-5)$$

$$\alpha(k) = rem \left\{ \frac{(C_{j_w} - Q_{j_w}(k)) \cdot l_{veh}}{N_{j_w} \cdot v_{j_w}(k) \cdot c(k)} \right\} \quad \beta(k) = floor \left\{ \frac{(C_{j_w} - Q_{j_w}(k)) \cdot l_{veh}}{N_{j_w} \cdot v_{j_w}(k) \cdot c(k)} \right\} \quad (4-6)$$

Where  $floor\{x\}$  = the largest integer that is smaller than or equal to  $\{x\}$

$rem\{x\}$  = the remainder

$c(k)$  = cycle length at time step  $k$

$l_{veh}$  = average vehicle spacing

The average approaching speed equals to free flow speed when speed limit is higher; otherwise, it equals to speed limit when speed limit is lower.

$$v_{j_w}(k) = \begin{cases} v_{j_w}^{free} & \text{under high speed limit} \\ v_{j_w}^{limit} & \text{under low speed limit} \end{cases} ;$$

### C. Merging into Lane Groups

Upon arriving at the end of a queue at a link, vehicles may change lanes and should merge into different lane groups, according to the driver's destination. The merging flow into lane group  $o$  at time step  $k$ , can be approximated:

$$q_{j_w,o}^{arrive}(k) = \beta_{j_w,o}(k) \cdot q_{j_w}^{arrive} \quad (4-7)$$

Where  $\beta_{j_w,o}(k)$  = the turning ratio for different turning movements. This study considers  $\beta_{j_w,o}(k)$  as predefined, and there is a large body of research on real-time O-D estimation.

#### **D. Departing Process**

The next step is the queue discharge for different lane groups  $o$ . The departing flow  $q_{j_w,d}^{out}(k)$  from different lane groups at time step  $k$  is given by:

$$q_{j_w,d}^{out}(k) = \min \left\{ S_{j_w,o}(k) \cdot g_{j_w,o}(k) / T, Q_{j_w,o}(k) / T + q_{j_w,o}^{arrive}(k), (C_{j_w,d} - n_{j_w,d}(k)) / T \right\} \quad (4-8)$$

Where  $S_{j_w,o}(k)$  = saturation flow rate of lane groups  $o$

$g_{j_w,o}(k)$  = green phase duration of lane groups  $o$

$Q_{j_w,o}$  = Queue length of lane group  $o$

$d$  = belongs to the set of downstream nodes of output links of link  $j_E$

$n_{j_w,d}(k)$  = number of vehicles in the link

The first term of Equation (4-8) considers the congested condition; the second term considers the uncongested condition; and the third term considers the available storage space of the destination link.

Saturation flow rate is calculated by using the method from HCM 2010, which estimates the saturation flow rate of any lane group based on known prevailing traffic parameters. The algorithm takes this form:

$$S_i = S_o \cdot N \cdot \prod_i f_i \quad (4-9)$$

Where  $S_o$  is the saturation flow rate per lane under base conditions and  $f_i$  is multiplicative adjustment factor for each prevailing condition  $i$ .

### E. Queue Evolution

Queues at lane groups are updated at every time step  $k$ .

$$Q_{j_w,o}(k+1) = Q_{j_w,o}(k) + T \cdot (q_{j_w,o}^{arrive}(k) - q_{j_w,d}^{out}(k)) \quad (4-10)$$

### F. Flow Conservation

The evolution of the total number of vehicles present at link  $j_w$  can be stated as:

$$n_{j_w}(k+1) = n_{j_w}(k) + (q_{j_w}^{in} - \sum_{d \in D_{j_w}} q_{j_w,d}^{out}) \cdot T + (d_{j_w} - e_{j_w}) \cdot T \quad (4-11)$$

$d_{j_w}$  and  $e_{j_w}$  are the demand flow and exit flow of links during time step  $k$ , respectively.

## 4.3 Optimization Formulation

### 4.3.1 Optimization Framework

This study adopts the MPC (Model Predictive Control) approach [92] to develop a model-based adaptive control strategy which addresses several issues aforementioned. Figure 4.3 illustrates the basic elements of the MPC control loop. The kernel of the control loop is the MPC controller, whose task is to specify, in real time, the control inputs to achieve the pre-specified objectives and constraints.

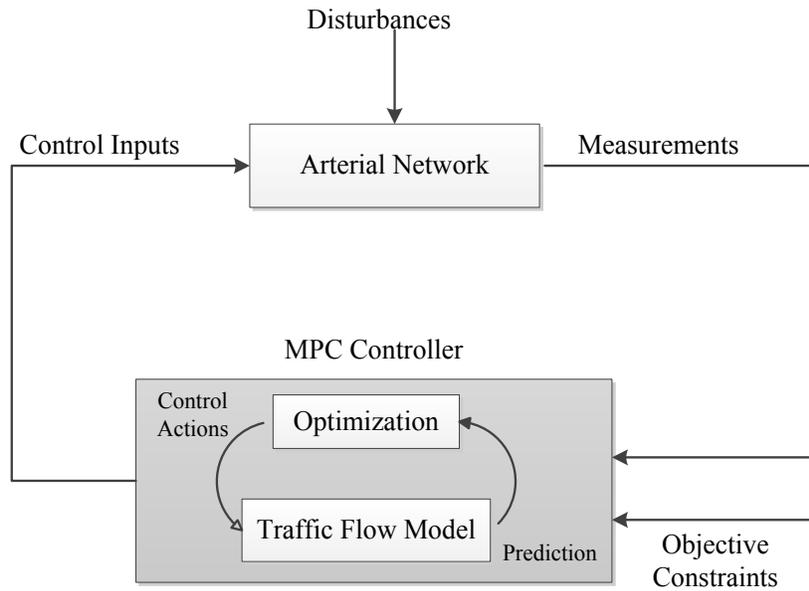


Figure 4.3 Framework of the Control Loop

In general, traffic flow models help adaptive control perform more proactively, although the models may also introduce errors that can be propagated (spatially and temporally) during the course of control actions. Many of the reviewed adaptive control systems adopt the rolling horizon procedure to overcome this problem. A modified rolling horizon scheme is used in this research. The concepts of control horizon and projection horizon keep the same, but they are time-variant in response to real-time traffic conditions. The following variable-time-window rolling horizon scheme is adopted in this study, as shown in Figure 4.4 [93].

- The stage, called as projection horizon, is the period over which traffic states are projected and the optimization problem is solved. It is integers of the optimized cycle length in that stage,  $S_p = M \cdot c(k)$ . The cycle length

is variable to adapt to the time-varying traffic conditions, as discussed in the abovementioned upper layer.

- Although one stage optimizes the control plan over  $S_p$  by using initial traffic measurements and demand predictions over  $S_p$ , but it is implemented only within the control interval  $T_k$ . After that, the projection and control horizon shifted forward by  $T_k$ . The optimization process starts again with collected new real-time measurements.

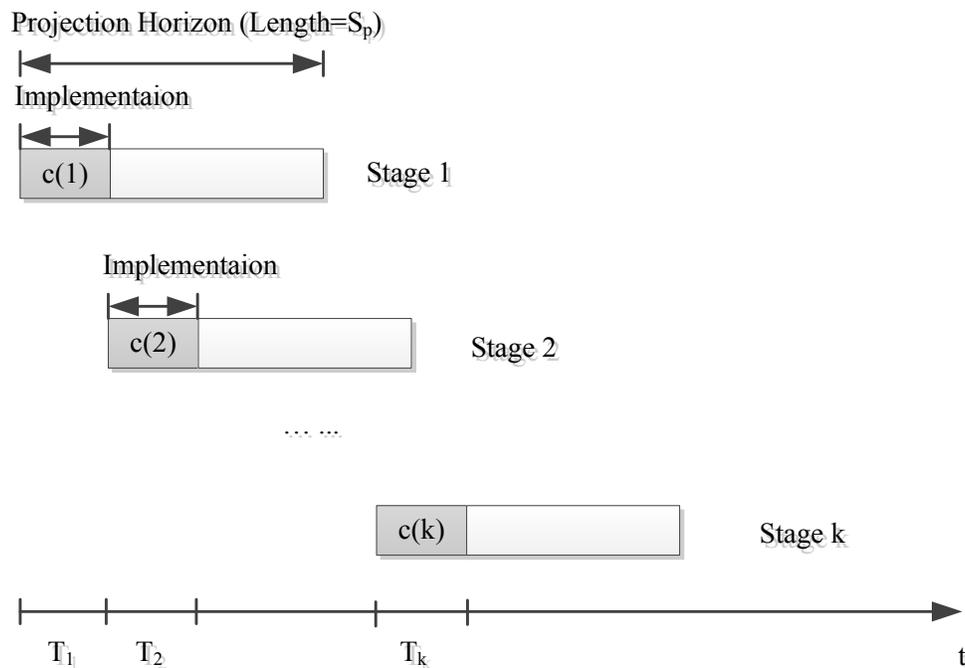


Figure 4.4 Illustration of the Rolling Horizon Scheme

### 4.3.2 Optimization Model

In congested conditions, the control objectives need to be decidedly different, as mobility is restricted. For example, the delay minimization strategy provides user-optimal delay minimization in uncongested conditions, but can sometimes work not in favor of minimizing total delay when systems become congested. Instead, the

signal plans should be timed such that every green second should be serving traffic at its maximum flow rate. In this research, the following represents the objective for maximizing the throughput in the controlled sub-network.

$$\max \sum_{k=1}^M \sum_{l \in L} q_l^{in}(k) \quad (4-12)$$

One type of the principal constraints is Equations (4-1), (4-5), (4-6), (4-7), (4-8), (4-9), (4-10), and (4-11), which represent the dynamic traffic state evolution along the arterial network. Another is the queue length constraints for left-run and through queues, as shown in Equation (4-13). The queue length cannot be larger than the capacity of the corresponding lane groups.

$$Q_{l,o}(k) \leq C_{l,o} \quad (4-13)$$

As the enhanced SFM considers different movements of one link, the two-ring, eight-phase structure from National Electrical Manufacturers Association (NEMA) is formulated as another type of constraints. The default phase mapping in the NEMA standard is shown in Figure 4-5 and Figure 4-6 [94].

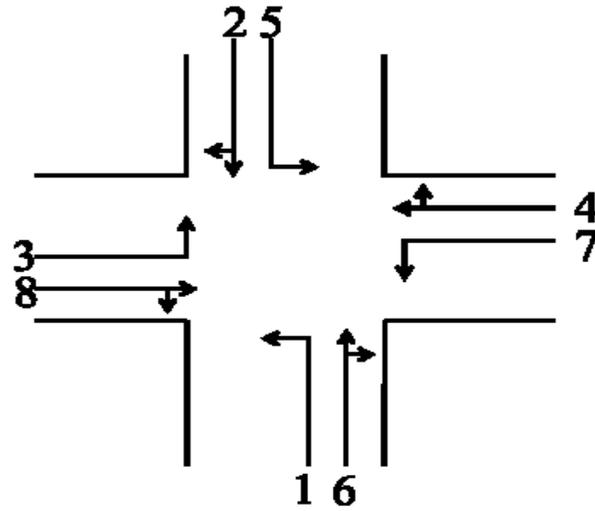


Figure 4.5 Typical Vehicular and Pedestrian Movements at a Four-leg Intersection

[94]

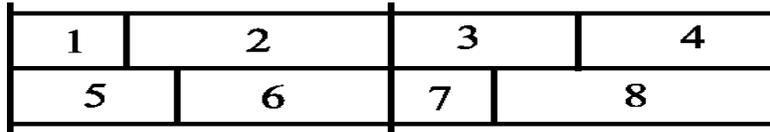


Figure 4.6 Standard Ring-and-barrier Diagram [94]

The decision variables are green durations. The constraints are composed of the physical structure of signal controllers and actual traffic conditions. In North America, the commonly accepted constraints are composed of three parts: 1) maximum and minimum greens; 2) pedestrian settings; and 3) cycle length and NEMA dual ring structure.

$$g_{j_N}^{left}(k) + g_{j_S}^{through}(k) = g_{j_W}^{left}(k) + g_{j_E}^{through}(k) \tag{4-14}$$

$$g_{j_S}^{left}(k) + g_{j_N}^{through}(k) = g_{j_E}^{left}(k) + g_{j_W}^{through}(k) \tag{4-15}$$

$$g_{j_N}^{left}(k) + g_{j_S}^{through}(k) + g_{j_W}^{left}(k) + g_{j_E}^{through}(k) = c(k) \tag{4-16}$$

Where  $g_{j_N}^{left}(k)$  and  $g_{j_N}^{through}(k)$  represent the green split for left turn and through movement of approach  $j_N$ , respectively;  $g_{j_S}^{left}(k)$  and  $g_{j_S}^{through}(k)$  represent the green split for left turn and through movement of approach  $j_S$ , respectively;  $g_{j_W}^{left}(k)$  and  $g_{j_W}^{through}(k)$  represent the green split for left turn and through movement of approach  $j_W$ , respectively;  $g_{j_E}^{left}(k)$  and  $g_{j_E}^{through}(k)$  represent the green split for left turn and through movement of approach  $j_E$ , respectively.

The following is the common minimum and maximum green constraint.

$$g_{l,o}^{\min} \leq g_{l,o}(k) \leq g_{l,o}^{\max} \quad l \in L \quad (4-17)$$

### 4.3.3 Solution Algorithm

The implementation of the GA is performed by a Genetic Algorithm Toolbox in MATLAB. The process of GA algorithm to search the optimal control action is shown in Figure 4.7. First, a population of solutions is generated, which represent a set of random possible control actions. Then the fitness of each chromosome is evaluated using the cost function and the constraints. Based on evaluation results, good chromosomes are randomly selected from the current population to be parents. Two new offspring chromosomes are produced from two parents according to a cross-over rule. The process of genetic operators repeats to produce new generation of possible solutions until reaching the stopping criteria. Finally, the best solution remains in the population is the final optimal control actions.

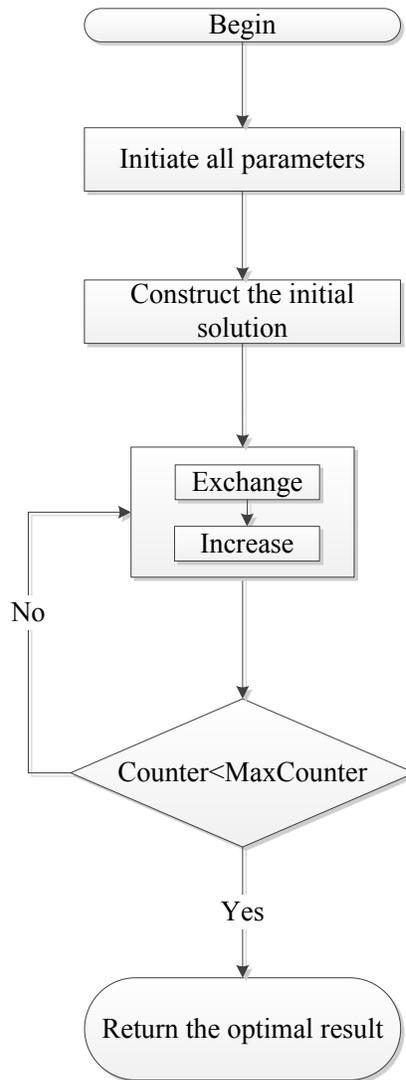


Figure 4.7 GA Process

#### 4.4 Simulation Evaluation

For comparison, VISSIM is employed as the performance index provider. The performance of the proposed model was compared with the benchmark, actuated control. ASC/3 treated each intersection independently and applied actuated control at each isolated intersection based on the optimal signal timing plans obtained from SYNCHRO 7.0. As this study only discuss the mid layer of split optimization, the

cycle length and offsets keep consistent during the simulation. Since the offset optimization is not implemented during the simulation experiments, it is not reasonable to use delay as a Measure of Effectiveness (MOE). Hence, the network throughput and maximum queue length are selected as MOEs. Due to the stochastic nature of the simulation model and underlying processes, each microsimulation run can be regarded as a random experiment, i.e. a random day in real life. Therefore 10 simulations with a common set of random seeds were completed for each scenario including the base case model. The final results, averaged over the multiple runs, were reported. When comparing different strategies, a representative run (median run) was then chosen.

Parameters used in the GA process are shown in Table 4.1. Figure 4.8 illustrates the convergence of GA results at each generation. The results represent the fitness function value of the best-fitted individual.

Table 4.1 Parameters Setting of GA

Name	Value
Population Size	20
Maximum trials for generating initial solutions	50
Number of offspring generated in each generation	100
Probability of carrying out local search (mutation)	0.3
Maximum number of generations	200

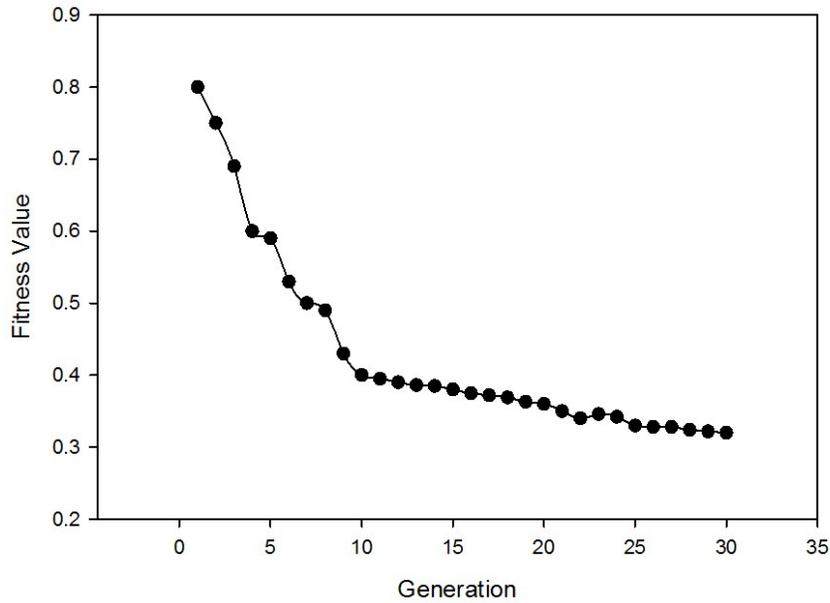


Figure 4.8 GA Results for Each Generation

Int. 1, Int. 2, Int. 5 and Int. 8 of downtown corridor have left-turn pocket lanes, and Int. 1, Int. 2, Int. 3, Int. 5 and Int. 6 of southeast corridor have left-turn pocket lanes. The left turn volume from the corridor to cross roads varies from 12 to 187 vehicles per hour. Table 4.2 presents comparison results from the proposed model and actuated control under different demand levels, based on one-hour simulations in VISSIM after a 10-min warm-up period. As indicated in Table 4.1, for the southeast corridor, the proposed model outperforms actuated control in average throughput by about 5% under high volume scenarios, and 2% under current volume scenarios, respectively. For the downtown corridor, the proposed model outperforms actuated control in average throughput by about 12% under high volume scenarios and 7% under current volume scenarios, respectively. The proposed model always

has the highest throughput compared with actuated control, especially under the higher demand level.

Another interesting finding is that the downtown corridor experienced greater improvements than the southeast corridor. This may be explained by the corridors' geographical configuration. The average spacing between two intersections are 120 meters and 720 meters for downtown and southeast corridor, respectively. It is reasonable that the proposed model performs better for the corridor with closely spaced signalized intersections because the uncongested part of the link is considered negligible compared to the total link length, and a platoon cannot be dispersed. The proposed traffic flow model well represents the stop-and-go traffic flow dynamic of signalized arterial network in congested conditions.

Table 4.2 Throughput Comparison of VISSIM Simulation Results

Corridor	Scenarios	MOE	Simulation Results from VISSIM		
			Proposed Model	Actuated	Improvement
Downtown Corridor	Current	vehicles in one hour	23882	22320	7%
	15%		27516	24568	12%
Southeast Corridor	Current	vehicles in one hour	19301	18923	2%
	15%		22646	21568	5%

The Box Plots of Figure 4.9 provide basic information about the distribution of different simulation runs. “1” represents the throughput distribution with actuated control under current demand; “2” represents the throughput distribution with proposed model control under current demand; “3” represents the throughput

distribution with actuated control under 15% increase demand; and “4” represents the throughput distribution with proposed model control under 15% increase demand. The throughputs from proposed model are remarkably higher than those from actuated control under two different demand conditions. Furthermore, the throughput spread from actuated control is much larger than that from proposed model control, pointing to a larger diversity of control performance. It seems the performance of the proposed model is much more stable, especially under the high demand condition. SYNCHRO selected longer cycle lengths to maximize the phase capacity for high demand scenario, this may adversely increase the chance of blockages due to the higher arrival rates to downstream links. Under actuated control scenarios with high demands, the severe blockages between lane groups and upstream-downstream links in the network can always be observed from VISSIM simulation animations. This phenomenon may explain the performance diversity of actuated control. In addition, we can also see that the downtown corridor experienced greater improvements than the southeast corridor.

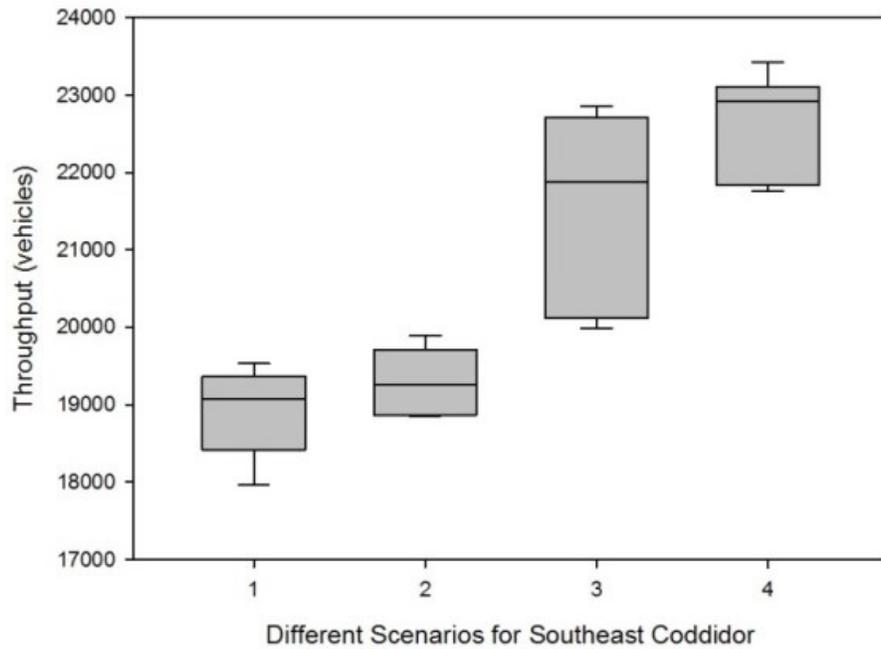
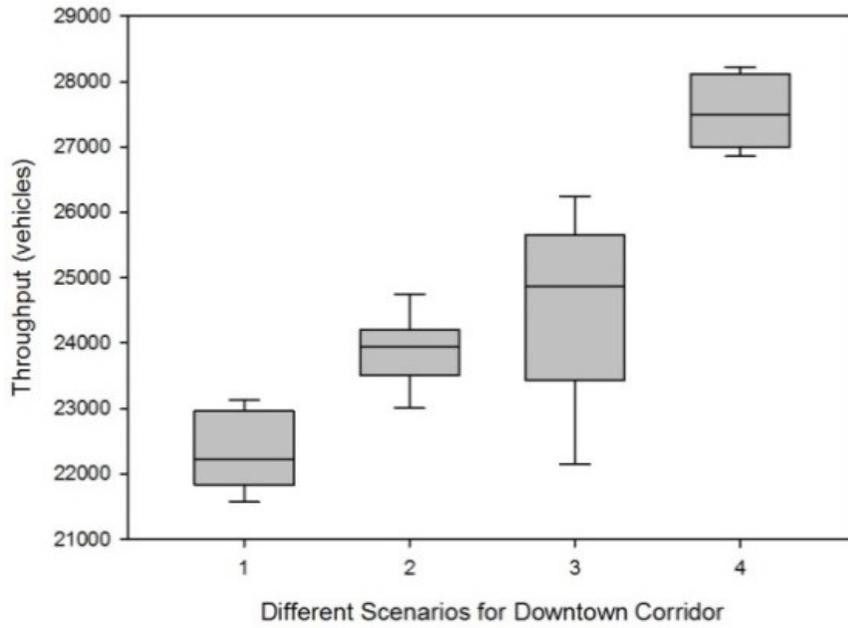
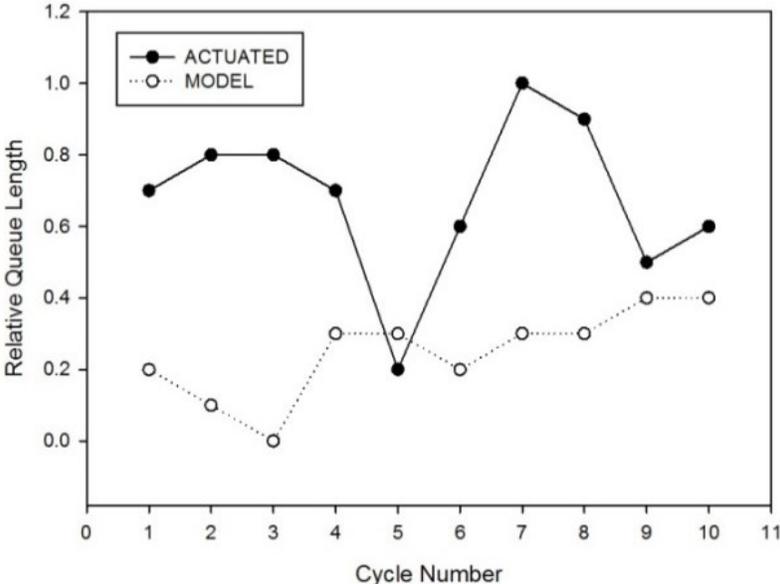


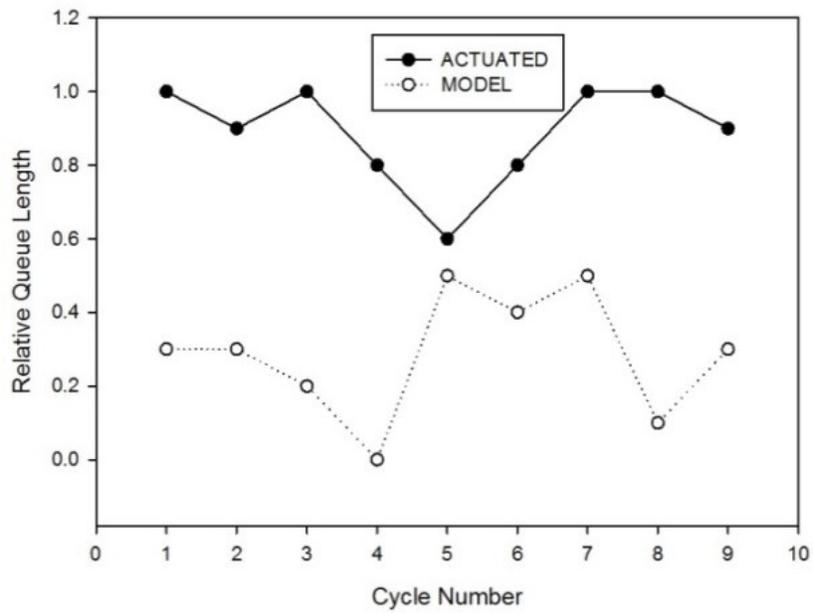
Figure 4.9 Distribution of Throughput under Different Scenarios

To test the capability of the proposed model with respect to capturing blockages between different lane groups and between downstream-upstream links

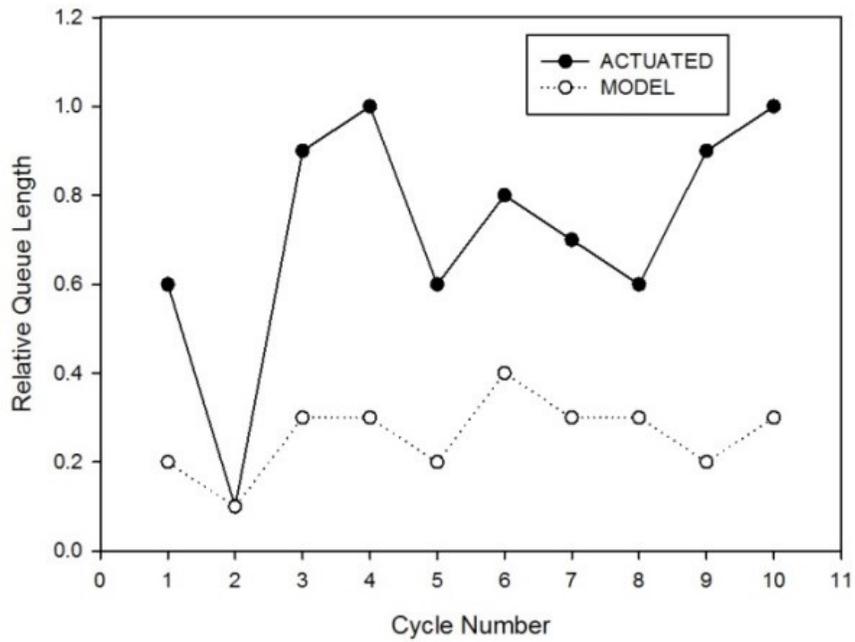
under congested conditions, the relative queue length ratio  $Q_i(k)/C_i(k)$  is used as the MOE. Ten consecutive values are analyzed from two particular intersections: 1) the through movement of westbound approach of Intersection 5 of the downtown corridor; and 2) the left-turn movement of southbound approach of Intersection 5 of the southeast corridor. As shown in Figure 4.10, the relative queue length ratio becomes higher when traffic demand is at 15% increase. Under the actuated control scenario, queues length frequently reaches the maximum length, whereas the proposed model yielded a smaller queue length and effectively mitigated the blockage and spillback. Furthermore, the proposed model can realize much more stable queue length for the consecutive period.



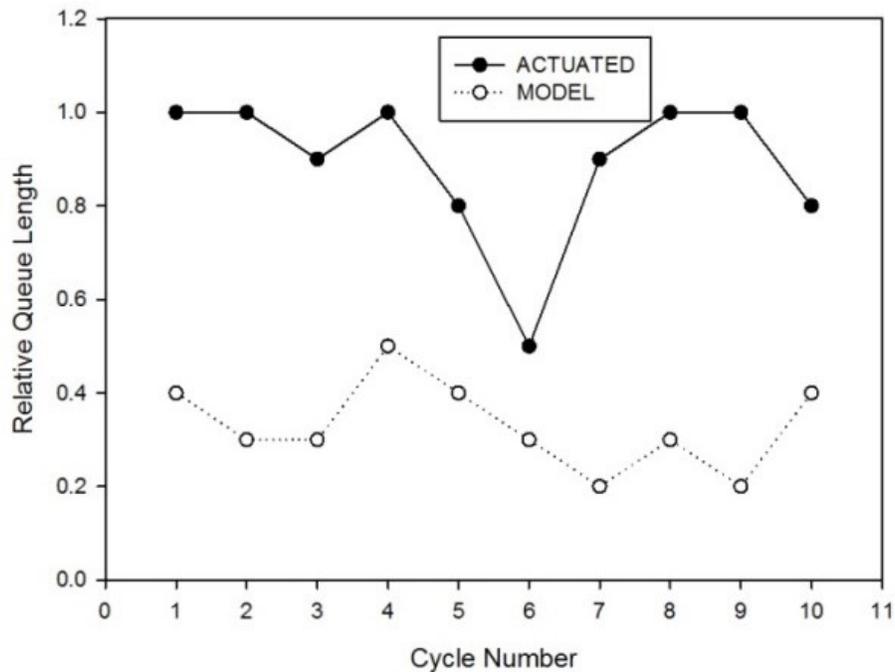
(a) Intersection 5 with Current Volume



(b) Intersection 5 with 15% Increase



(c) Intersection 4 with Current Volume



(d) Intersection 4 with 15% Increase

Figure 4.10 Relative Queue Length Dynamic.

#### 4.5 Summary

This chapter presented an optimization model for real-time signal control under congested conditions. The proposed model combines an enhanced SFM, mathematical optimization and optimal control, to capture not only the critical operational issues at signalized intersections, but also the traffic evolution along the arterial link as well as to ensure computing efficiency. Clearly, SFM is a simple model that can only provide a rough representation of the traffic dynamic in oversaturated conditions. However, SFM represents the stop-and-go traffic flow dynamic of signalized arterial network fairly well in congested conditions. As the model is only applicable for split optimization, a hierarchical control structure was proposed to optimize the cycle length, split and offset. The highest layer updates the

cycle length over time based on network capacities and volume levels. The mid layer continuously calculates optimal split using the rolling horizon scheme. Finally, the offsets at each intersection are optimized.

For the split design under two different traffic demand patterns, the presented microscopic simulation shows the effectiveness of the proposed model in comparison with actuated control based on the optimal signal timing plans obtained from SYNCHRO. The throughput improvement ranges from 2% to 12% depending on the demand patterns. One interesting finding is that geometry configurations may affect the performance of the model. It is reasonable that the traffic model performs better for the corridor with closely spaced signalized intersections, because the uncongested part of a link is considered negligible compared to the total link length, and a platoon cannot be dispersed. From the results of queue length dynamic, the proposed model successfully manages the queue length to avoid spillback and blockage. This extensive simulation experiment and analyses in comparison with results from SYNCHRO reveals that the proposed model is promising for use in the design of arterial signals, especially under congested, high demand traffic conditions.

# **Chapter 5 Adaptive Model-based Offsets Optimization for Congested Arterial Traffic**

## **5.1 Introduction**

Proper determination of intersection offsets provides for the efficient movement of platoons through multiple intersections during the green signal phase, resulting in significantly reduced delays and improved driver satisfaction. In past research, two major strategies for developing signal coordination timing plans have been used: bandwidth maximization [7, 95-98] and flow profile methods [68, 99-101]. These two methods optimize offsets according to several mathematical objectives, such as maximizing bandwidth [7, 96] and minimizing disutility (e.g. delay or number of stops) [102]. A number of researchers have also proposed various approaches to improve the performance of actuated coordinated systems [103-106]. However, the performance of these signal coordination algorithms may lead to suboptimal results during the congested condition, which is characterized by a lack of steady-state flow conditions and by significant interaction among traffic state between adjacent links. Instead of based on the average traffic flow conditions and predefined congestion conditions, dynamic offset optimization is needed to deal with the congestion phenomena, such as dynamic evolution of queues at intersection approaches.

Adaptive Traffic Control Systems (ATCS) optimize traffic signal control in real-time by continuously adapting signal timing plans at intersections to the current traffic demand [10]. With advances in computation and sensing, ATCS have become

an increasingly attractive research topic and traffic control option. A number of elaborate traffic flow models for the signalized arterial network have been deductively derived to describe the complex interactions between traffic states evolution and key signal control parameters [14, 34, 39, 64, 90]. Optimization tools are used to search for the best sequence of control decisions based on the traffic information predicted by traffic flow models. Hence, the undesirable traffic conditions will not develop by employing embedded traffic flow models. For some model-based ATCS, the signal timing plans, including the cycle length, green split, offset and phase sequence, are optimized in a centralized manner simultaneously. The problem can be described as a mathematical optimization problem and the objective is to minimize or maximize a performance measure that is a complex function of the traffic state and signal timing parameters. However, the implementation is constrained by the real-time computational complexity of the NP-hard optimization problem [34, 77]. Some other ATCS develops hierarchical structure to divide the complex control problem of a large traffic system into different control levels or layers. Control problems with different details are addressed in different levels, e.g. the coordination layer optimizes the offsets at each intersection (one per cycle), and local control layer continuously calculates optimal splits.

In this study, the offset is optimized based on a developed hierarchical structure of signal timing optimization. In the highest layer, the cycle length is adjusted based on the predicted traffic demand and network saturation rate. Over

time, the cycle length is updated as the system adapts to changing traffic conditions. In the mid layer, the split is optimized proactively with an embedded traffic flow model and in a rolling horizon scheme to maximize the throughput. Through appropriate state equations, the control procedure ensures time-dependent, dynamic control. Lastly, the offset is adjusted based on the optimized cycle length and green split. This section will discuss the offset optimization layer. This study expands the dynamic queue concept to the network-wide coordination problem. First, the ideal offset are found based on the shockwave profiles at each signalized intersection. Then, goal programming is introduced to optimize offset for the whole network. Simulation experiments are conducted to compare the proposed model with fixed-time control in producing network-wide coordination.

## **5.2 Methodology**

### **5.2.1 Ideal Offset and Boundary Offsets**

Using a shockwave theory, numerous studies have developed queue dynamic processes for signalized intersections under congested conditions [74, 79, 107, 108]. According to the Lighthill-Whitham-Richards (LWR) theory [109], shockwaves are generated by the traffic signal at intersections. Figure 5.1 is a space-time diagram to display queue dynamics under congested conditions. At the beginning of the effective green phase, the front of a residual queue begins to discharge at the saturation flow rate and a discharge shockwave propagates upstream from the stop line of intersection  $i$ . The platoon from intersection  $i-1$  entering the link  $(i-1, i)$  encounters the residual queue  $Q_{(i-1, i)}(k)$  at intersection  $i$  at time step  $k$ . A backward-

moving shockwave is created by the stoppage caused by the residual queue. Therefore, subsequent entering vehicles encounter stoppage. Whether these two shockwaves intersect with each other depends on the relationship between the saturated discharging traffic flow at intersection  $i$  and the traffic arrival from intersection  $i - 1$ . A new residual queue is formed sometime after the start of the red light of the next cycle when the queuing shockwave meets the traffic arrival. The shockwaves and queue dynamic described above will repeat from cycle to cycle.

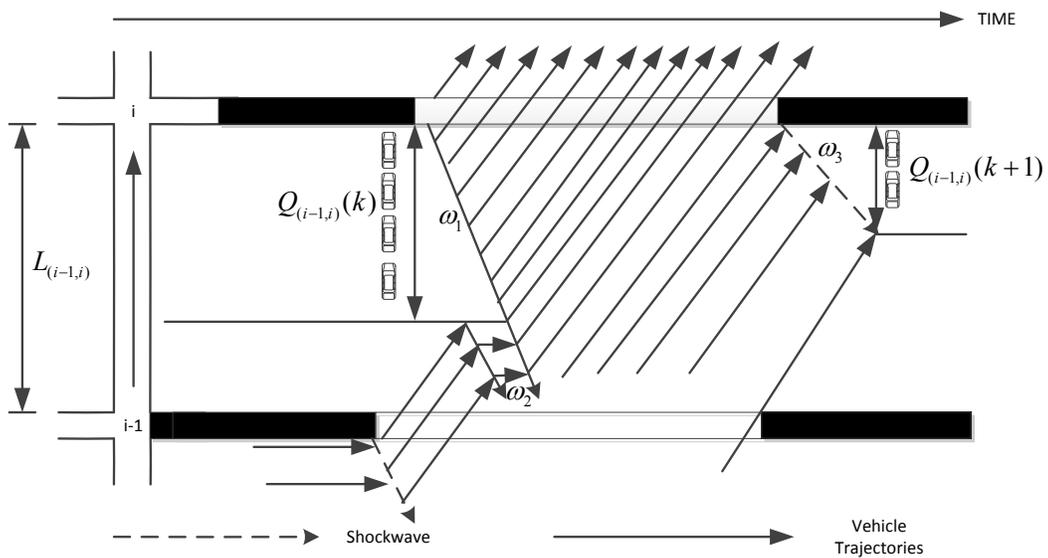


Figure 5.1 Shockwave Profile and Queue Dynamic in Congested Condition

From Figure 5.1, it is observed that if the discharged platoon from intersection  $i - 1$  joins the tail of the downstream residual queue at the time when the tail has reached its free flow speed, then there will be no stoppage or starvation. This ideal signal offset allows the leading vehicle in the incoming platoon to just avoid encountering the residual queue, yet allows it to reach the stop line one headway after the last vehicle in the residual queue discharges. This ideal offset was

calculated via several different equations in previous research, such as [79] and [67]. In this study, Equation (1) by Lieberman et al. [79] is used to calculate the ideal offset.

$$\phi_{(i-1,i)}^{ideal}(k) = \frac{L_{(i-1,i)}}{v_{(i-1,i)}} - \left[ \frac{(v_{(i-1,i)} + \omega_1)}{v_{(i-1,i)}\omega_1} \right] Q_{(i-1,i)}(k) \quad (5-1)$$

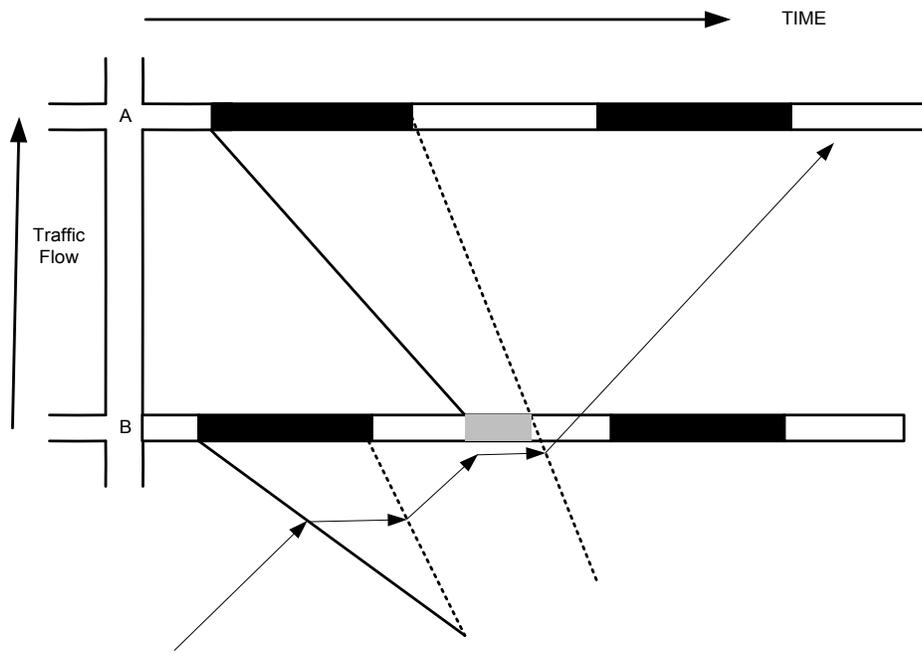
Where  $L_{(i-1,i)}$  = the link length

$v_{(i-1,i)}$  = the travel speed of the leading vehicle of the incoming platoon

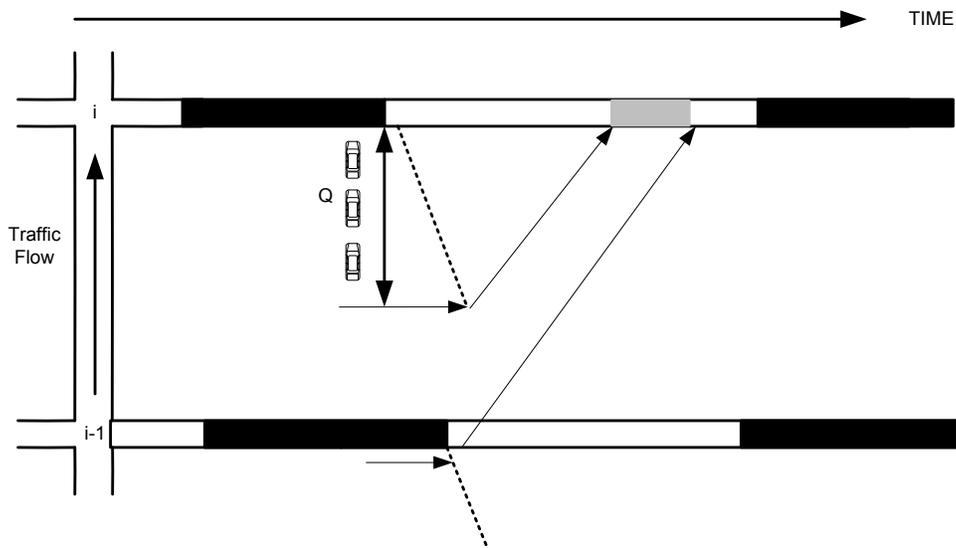
$\omega_1$  = discharge shockwave speed

$Q_{(i-1,i)}(k)$  = residual queue length

As shown in Figure 5.2 (a), spillback occurs when a queue from a downstream intersection uses up all the space on a link and prevents vehicles from entering the upstream link on green. Some literature has also defined this condition as causing “de facto red” to the upstream movement. As shown in Figure (b), starvation occurs at a downstream intersection when the discharge of vehicles at the upstream intersection is delayed beyond the ideal offset. Starvation causes capacity loss due to the wasting of limited green time at the downstream intersection. These two phenomena both waste green time and should be avoided.



(a) Spillback



(b) Starvation

Figure 5.2 Spillback and Starvation in Congested Condition.

This study defines two boundary offsets: maximum offset and minimum offset. The maximum offset prevents spillback at upstream intersections by ensuring that the stoppage shockwave  $\omega_2$  dissipates before reaching the upstream intersection. The minimum offset ensures that the first-released vehicle joins the discharge queue at the downstream intersection. Equations (2) and (3) by Lieberman et al. [79] are used to exemplify the concept.

$$\text{Maximum Offset} = \frac{L_{(i-1,i)}}{v_{(i-1,i)}} \left[ 1 - \frac{Q_{(i-1,i)}}{L_{(i-1,i)}} \left( 1 + \frac{v_{(i-1,i)}}{\omega_2} \right) \right] + \min \left[ g_{(i-1,i)}, \frac{L_{(i-1,i)}}{\omega_2} \right] \left( 1 - \frac{\omega_2}{\omega_1} \right) \quad (5-2)$$

$$\text{Minimum Offset} = \frac{L_{(i-1,i)}}{v_{(i-1,i)}} \left[ 1 - \frac{Q_{(i-1,i)} h v_{(i-1,i)}}{L_{(i-1,i)} l_{veh}} \right] \quad (5-3)$$

Where  $g_{(i-1,i)}$  = the green phase duration

$h$  = mean queue discharge headway

$l_{veh}$  = average vehicle spacing within a standing queue

$\omega_2$  = speed of backward-moving shockwave caused by stoppage

Another constraint is the offsets relationship between primary and opposing traffic. Taking the relationship between two offsets of one link as an example, two offsets are defined: the primary direction,  $\phi_{(i-1,i)}(m)$  and the opposing direction,  $\phi_{(i,i-1)}(m)$ . There are two different scenarios when this relationship is considered, which depends on the value of  $\phi_{(i-1,i)}(m)$ , as shown in Figure 5.3. If  $\phi_{(i-1,i)}(m) < 0$ , the offset is set to clear heavy queues for the primary directions and the opposing traffic tends to arrive at the same cycle as it is released from signal  $i$ . If  $\phi_{(i-1,i)}(m) > 0$ , the queue in the primary direction is lighter, and the green start time at intersection

$i-1$  is leading the green start time at intersection  $i$ . Then, Equation (7) describes the offset relationship between the two directions, where  $m$  is the index for a cycle number,  $c(k)$  is the cycle length for signal  $i-1$  at cycle  $k$ , and  $n$  is the number of cycles of signal  $i-1$ , reflecting the traffic regimes of the two intersections.

$$\begin{cases} \phi_{(i-1,i)}(k) + \phi_{(i,i-1)}(k) = 0 & \text{If } \phi_{(i-1,i)}(k) < 0 \\ \phi_{(i-1,i)}(k) + \phi_{(i,i-1)}(k) = \sum_{m=k}^{k+n} c_{i-1}(m) & \text{If } \phi_{(i-1,i)}(k) > 0 \end{cases} \quad (5-4)$$

Where  $m$  = an integer value, reflecting the traffic regimes of the two intersections.

The offsets of another two directions also need to satisfy this relationship. In addition, this method does not support half-cycling.

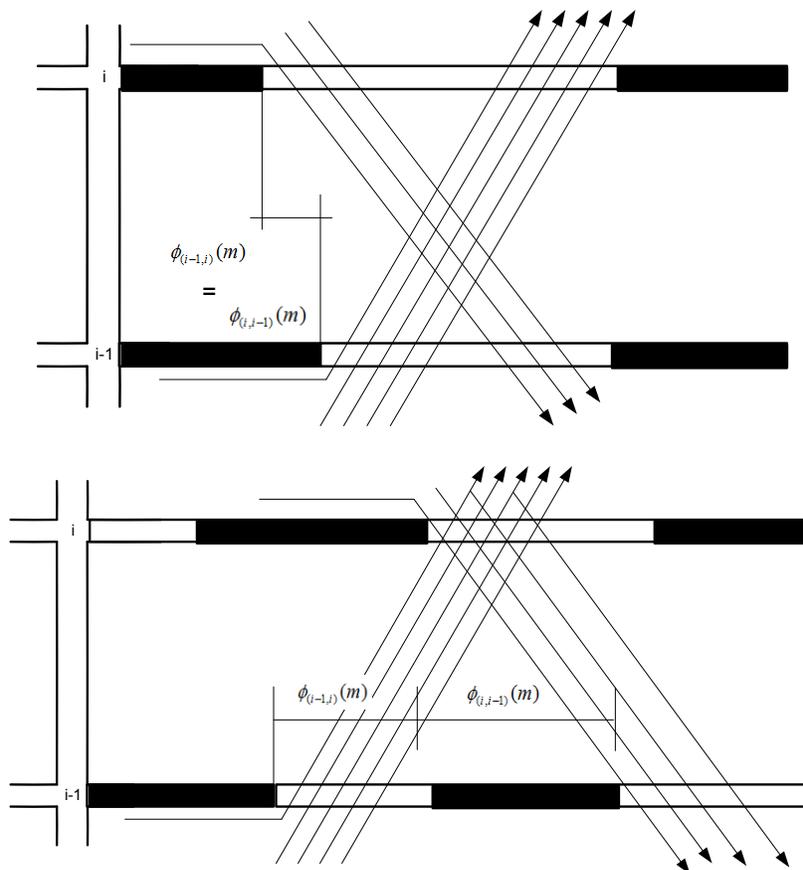


Figure 5.3 Relationships between Offsets for Primary and Opposing Traffic

### **5.2.2 Model Formulation**

To use the available green phase duration and minimize delays, the offset needs to be intuitively designed to control the interaction between incoming platoons and residual queues. This depends on the input-output flow balance and queue length control in each congested approach at every cycle. As aforementioned, the highest layer and mid layer of the hierarchical structure adjusted the cycle length and split to control the input-output flow balance and queue length dynamic; now, the question is how to optimize the offset from the network level. As discussed previously, if the offset exceeds the maximum and minimum offsets, it could result in a condition of spillback and starvation. Therefore, it is essential to control the offset within this certain range. This study adopts the approach of goal programming to formulate this problem in order to provide a compromising solution among multiple objectives. Generally, the modeling process of goal programming approach is: 1) to establish a specific numeric goal for each of the objectives, 2) to formulate an objective function for each objective, and 3) to seek a solution that minimizes the sum of deviations of these objective functions [110]. There are different types of goal programming approach according to how goals compare in importance. Specifically, the approach of preemptive goal programming is used in this study to find a solution of the offsets at intersections, which has a hierarchy of priority levels for different goals. In this case, first-priority consideration is the primarily important goal; second-priority consideration is the secondarily important goal; and so forth. According to the relative importance of single control objectives, the optimization problem is

formulated 1) to minimize spillback, and 2) to minimize the starvation, 3) to maintain the ideal offset. The first objective is deemed the most important for signal control under congested conditions because it can avoid intersection blockages. Then, the first objective is converted to goal constraints as upper one-sided goals that we do not want to exceed. This is because the smaller the objective values, the better the offset control can manage spillback. Each goal can then be expressed as follows:

$$\text{Goal 1: } f_1 = \min (\phi_{(i,i-1)}(k)) \quad (5-5)$$

$$\text{Goal 2: } f_2 = \max (\phi_{(i,i-1)}(k)) \quad (5-6)$$

$$\text{Goal 3: } f_3 = \min \left\{ \begin{array}{l} \sum_{l \in L} q_{(i-1,i)}^{in}(k) \left( \phi_{(i-1,i)}(k) - \frac{L_{(i-1,i)}}{v_{(i-1,i)}} + \left[ \frac{(v_{(i-1,i)} + \omega_1)}{v_{(i-1,i)} \omega_1} \right] Q_{(i-1,i)}(k) \right) + \\ \sum_{l \in L} q_{(i,i-1)}^{in}(k) \left( \phi_{(i,i-1)}(k) - \frac{L_{(i,i-1)}}{v_{(i,i-1)}} + \left[ \frac{(v_{(i,i-1)} + \omega_1)}{v_{(i,i-1)} \omega_1} \right] Q_{(i,i-1)}(k) \right) \end{array} \right\} \quad (5-7)$$

Where  $f_l$  is the spillover goal,  $f_2$  is the starvation goal,  $f_2$  is the ideal offset goal,  $q_{(i-1,i)}^{in}(k)$  is the upstream arrival flows at time step  $k$  of link  $l$ .

Then, we convert goals to a preemptive goal program as follows:

$$\begin{array}{l} \text{Lexmin} \{ \rho_1^+, \rho_2^-, \rho_3^+ \} \\ \text{Subject to} \\ f_1 + \rho_1^+ - \rho_1^- = b_1 \\ f_2 + \rho_2^+ - \rho_2^- = b_2 \\ f_3 + \rho_3^+ - \rho_3^- = b_3 \end{array} \quad (5-8)$$

Where lexmin represents lexicographic minimization,  $b_i$  represents an aspiration level of  $f_i$ ,  $\rho_i^+$  represents a positive deviation from the aspiration level of  $f_i$ ,

and  $\rho_i^-$  represents a negative deviation from the aspiration level of  $f_i$ . In this research,  $b_1$ , which is the aspiration level of the most important objective, equals to the maximum offset,  $b_2$  equals to the minimum offset under which we do not want to fall, and  $b_3$  equals to zero. For a two-way arterial network, the ideal offset of each approach of an intersection cannot be achieved simultaneously, because they interact with each other.

### 5.2.3 Solution Algorithm

A standard GA cannot be applied directly to solve the above formulated problem which has three objectives with different relative priority. A systematic way proposed by Sherali and Soyster is used to convert the preemptive goal program to a single-objective optimization problem by adding a set of equivalent weights for each objective [111, 112]. Then the standard GA can be implemented to solve the converted problem. According to importance, the first objective dominates the second objective, so the added weights must also guarantee that the first objective always dominates the second objective. The proposed methodology proposed by Sherali and Soyster is as follows.

$$\begin{aligned} & \text{Minimize } \{f_1(x), f_2(x), f_3(x); f_1(x) \gg \gg \\ & \text{s.t. } x \in X \end{aligned} \tag{5-9}$$

Assume that  $x^*$  is the set of optimal solutions to (5-8). Then  $x^*$  is also an optimal solution to the following program:

$$\text{Minimize } \{F(x) = \lambda_1 \cdot f_1(x) + \lambda_2 \cdot f_2(x) + \lambda_3 \cdot f_3(x); x \in X\}$$

Where

$$\begin{aligned}
\lambda_i &= M^{(3-i)}; i=1,2,3 \\
M &\geq M_0 \\
M_0 &= 1 + \gamma \\
\gamma &= \max_i \{ \max(f_i) - \min(f_i) \}
\end{aligned} \tag{5-9}$$

The first priority goal (i=1) receives the highest weight. The second and third goals receive smaller weights. Finally, the preemptive goal program of (5-8) is transformed to a single-objective optimization formulation, which is written as

$$\begin{aligned}
&\text{Minimize } (M^2) \cdot (\rho_1^+) + (M) \cdot (\rho_2^+) + (\rho_3^+) \\
&\text{Subject to} \\
&f_1 + \rho_1^+ - \rho_1^- = b_1 \\
&f_2 + \rho_2^+ - \rho_2^- = b_2 \\
&f_3 + \rho_3^+ - \rho_3^- = b_3
\end{aligned} \tag{5-10}$$

The standard GA can be applied to solve the problem (5-10). The implementation of the GA is performed by a Genetic Algorithm Toolbox in MATLAB. Similar to Figure 4.7, the GA repeatedly modifies a population of individual solutions using selection rules, crossover rules, and mutation rules.

### 5.3 Simulation Evaluation

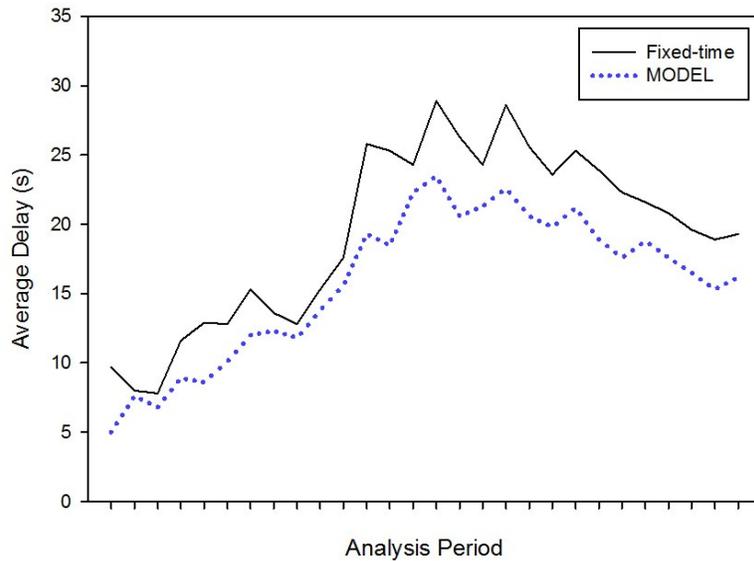
A reference case was needed for comparison with the proposed model. Therefore, one optimized fixed-time control plan was generated with SYNCHRO for both the current and 15% increase traffic demand. For the reference case, the offsets were constant during the VISSIM simulation. Each of the scenarios was simulated multiple times and results were tested for statistical significance. As this study only discusses the lower layer of offset optimization, the cycle length and split were

consistent during the simulation. In addition, the proposed model does not consider the transition. The scenario with 15% increase traffic demand exhibited extensive queues, which propagate to block the upstream intersection.

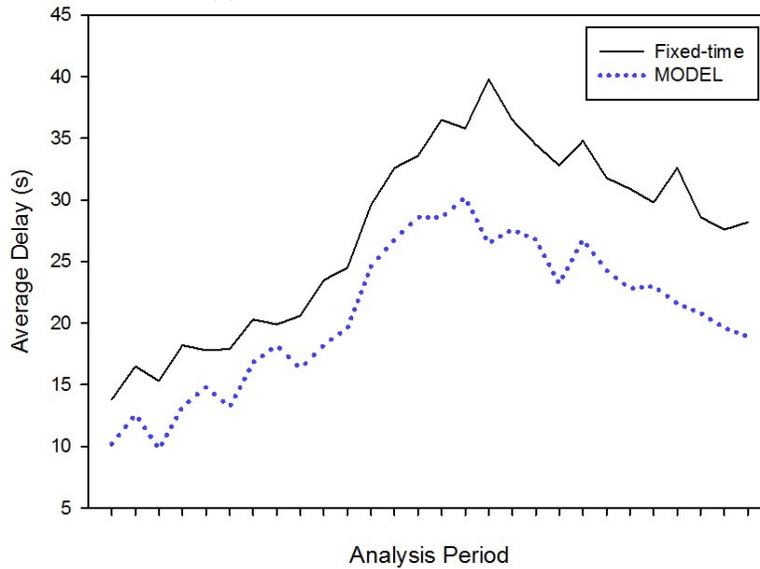
The average delay for the whole corridor was used as the Measures of Effectiveness (MOE). As shown in Table 2, the total average delay is improved at 9% and 14% under current and 15% increase demand scenarios, respectively. Figure 5.4 shows the average delays at different analysis periods. One can find out that, in general, the optimized offset generates shorter travel delay than the fixed offset. Therefore, the proposed model performs better in producing corridor-wide coordination in terms of total delay.

Table 5.1 Delay Comparison of VISSIM Simulation Results

Scenarios	MOE	Simulation Results from VISSIM		
		Proposed Model	Fixed-time	Improvement
Current	Average	19.1	20.9	9%
15%	Delay(s)	25.6	29.8	14%



(a) Current-Demand Scenario



(b) 15%-Increase Scenario

Figure 5.4 Average Delay under Different Demand Scenario

Another performance measure is the available space for queues in corridor's links. Low percentage values indicate higher chances of queue backup and possible spillback. Figure 5.5 shows the reserve queuing capacities percentages on the corridor links as time progress. Results indicate the effectiveness of the developed control plans in allocated queues in the corridor's links.

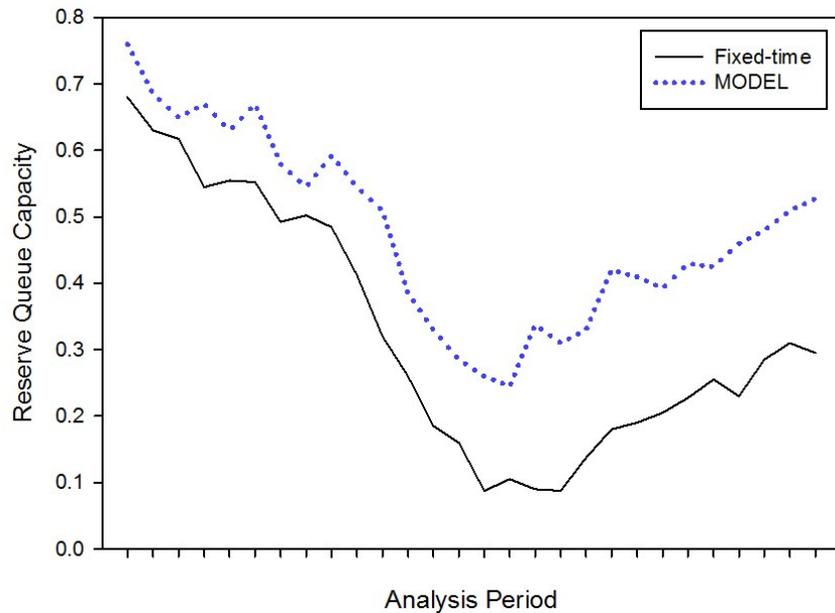


Figure 5.5 Network Reserve Queuing Capacities

#### 5.4 Summary

This study presented an algorithm to design signal coordination for networks with congested intersections. The ideal offset, maximum offset, and minimum offset were introduced to formulate the optimization function. Then, goal programming was introduced to optimize offset for the whole corridor. Simulation results showed that, in terms of total delay, the proposed model provided better coordination than fixed-time control plan did. This study only discussed offset optimization with a fixed cycle length and split. In the future, the whole architecture will be implemented in one optimization process to verify the performance. Another future study is to compare the performance of the proposed model with that of other adaptive offset optimization methods.

# **Chapter 6 Adaptive Signal Priority Control on Mixed Traffic Arterials**

## **6.1 Introduction**

The adaptive priority control is a typical multi-objective optimization problem, where two or more objectives must be satisfied simultaneously in order to obtain the preferred solution. Normally, objectives from different travel modes may be in conflict with each other. Each trade-off solution corresponds to a specific order of importance of the objectives from different travel modes. Various studies have used the preference-based method to balance the trade-offs between different travel modes, where the integrated delay is formulated as the objective of their control algorithms. In the preference-based method (i.e. weighted-sum method), a set of objectives are normalized and scaled into a single composite objective by summing the weighted normalized objectives. Then the task is to find one solution which optimizes the single-objective problem. Table 6.1 lists the objective functions of some existing adaptive priority control problem. The objective functions of most reviewed studies are to minimize the weighted summation of delays of various travel modes, although they used different delay estimation models. The weights can be considered as a function of various factors such as real-time bus occupancy, schedule lateness, maximum allowed traffic delay, longest queues and priority for emergency vehicles. The weighting factor is a relative number among various objectives, and represents system manager's relative preference on each objective. For example, the relative

weighting factor on bus delay over vehicular delay and pedestrian delay means the policy preference to shift more passengers from private vehicles to public transportation.

The drawback of the preference-based approach is obvious. First, the solutions are sensitive to the relative preference vector which is determined without any knowledge of the possible consequences. A change in this preference vector will result in a different solution. Second, finding a relative preference vector may be highly subjective, which sometimes requires experience-driven and qualitative information. In reality, some factors are often difficult to determine beforehand and hard to be weighted. Third, without perfect knowledge of the trade-off information between different objectives, the obtained solutions are circumstance dependent. Instead of using ambiguous weighting factors, the other approach is to generate a set of solutions, called Pareto-optimal solution, to investigate the trade-off information between different objective using multi-objective optimization technologies [113]. Then the decision can be taken after the optimization using high-level information. This study implements multi-objective optimization technology to evaluate the trade-off among different conflicting adaptive priority control objectives under mixed arterial traffic. Transportation manager will be in a better position to make a choice to balance the priority request from different travel modes when such trade-off solutions are unveiled.

Table 6.1 Summary of Preference-based Adaptive TSP Methods

<b>Authors</b>	<b>Performance Index</b>	<b>Model and Algorithm</b>
Li et al. [21]	A weighted sum of bus and other traffic delay	Standard quadratic programming models
Christofa and Skabardonis [23]	A summation of the person delay for the auto and transit vehicles passengers	Quadratic programming models
Christofa et al. [114]	A person based delay considering schedule adherence and auto vehicle progression	A mixed-integer linear program
Duerr [115]	A weighted sum of vehicle delay, vehicle stops, residual queues and overflow impact	Genetic Algorithm approach
He et al. [17, 19]	A weighted sum of signal delay, queue delay, platoon penalty, and the sum of slack variables	A mixed-integer linear program
Stevanovic et al. [116]	A summation of the person delay for the auto and transit vehicles passengers	Genetic Algorithm approach
Ma et al. [117]	Total weighted delay of all bus priority requests considering schedule deviation	Dynamic programming model, rolling time horizon approach
Lee et al. [118]	A weighted general traffic delay and transit vehicle delay	Genetic Algorithm approach
Medina et al. [119]	Delay based on unit, delay based on occupancy, delay based on occupancy and priority	Simple additive weighting, analytical hierarchical process, technique for order preference by similarity to ideal solution

## 6.2 Model Development

The multi-objective optimization procedure is shown in Figure 6.1. After all priority control objectives from different travel modes and control constraints have been defined, the multi-objective optimization technology will find a well-distributed set of trade-off solutions, which is called the Pareto optimal set. It means no improvement can be achieved in any objective without degradation in the other. A solution on the Pareto front is corresponding to a given objective weight set. The vectors of the decision variables corresponding to the solutions included in the Pareto optimal set are called non-dominated [113]. The non-optimal solutions are in the area above the Pareto front, and infeasible solutions are in the area below the Pareto front. Transportation manager can obtain valuable information based on the shape of the Pareto front. For example, they can know how much other objective functions would be compromised if a selected objective function is to be favored. Finally, one solution can be selected with high-level traffic information.

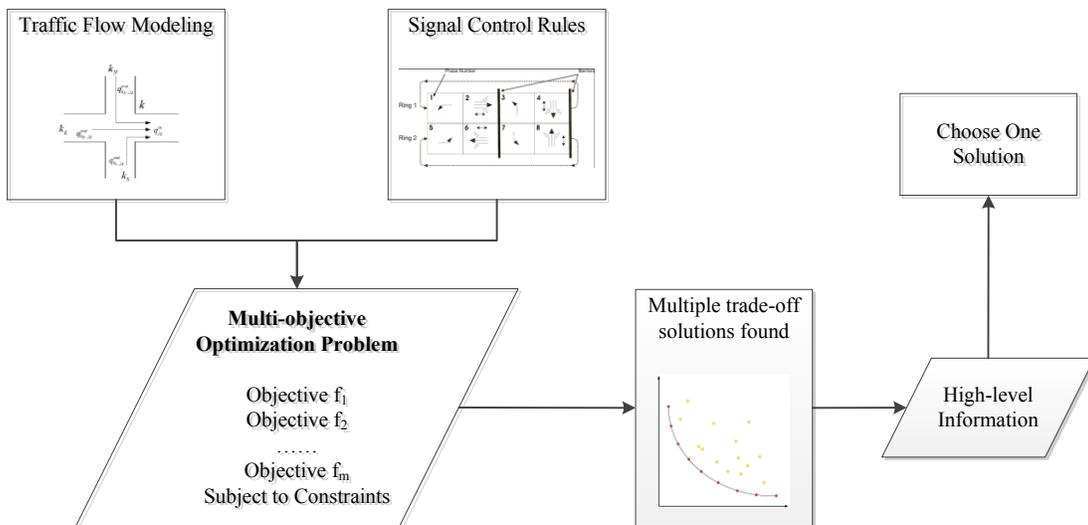


Figure 6.1 Multi-objective Optimization Process

### 6.2.1 Problem Formulation

The multi-objective adaptive priority control model is summarized as follows.

$$\text{Objective: Minimize } D(g) = |\max(d_a(g), \sum_N d_b(g))| \quad (6-1)$$

$$\text{subject to: } \left\{ \begin{array}{l} g_i - \max(\lambda(g_{walk} + g_{pedclearance}), g_i^{\max}) \leq 0 \\ -g_i + \min(\lambda(g_{walk} + g_{pedclearance}), g_i^{\min}) \leq 0 \\ \sum_{i=1}^4 (g_i + y + ar) - C = 0 \\ \sum_{j=5}^8 (g_j + y + ar) - C = 0 \\ g_1 + g_2 - g_5 - g_6 = 0 \\ g_3 + g_4 - g_7 - g_8 = 0 \end{array} \right. \quad (6-2)$$

Where  $g_i$  is green duration time of phase  $i$ ;  $g_i^{\min}$  is minimum green;  $g_i^{\max}$  is the maximum green;  $g_{walk}$  is walk time;  $g_{pedclearance}$  is pedestrian clearance time;  $\lambda$  is a flag variable (0: no pedestrian call; 1: pedestrian call);  $C$  is cycle length;  $y$  is yellow time; and  $ar$  is All-red time.

The variables in the optimization are green duration  $g_i$ . The first objective is to minimize the maximum control delay  $d_a$ , and the second objective is to minimize total bus delay  $\beta \sum_N d_b$ . Based on the HCM 2010, the control delay consists of two parts: uniform delay (UD) and random delay (RD), or uniform delay (UD<sub>o</sub>) and overflow delay (OD) under overflow condition [120].

If the volume to capacity ratio ( $X$ ) is smaller than 1, then

$$d_a = 0.9(UD + RD) = \frac{1}{2}C \frac{(1 - g / C)^2}{1 - (g / C)\min(X, 1)} + \frac{1}{2v} \left( \frac{X^2}{1 - X} \right) \quad (6-3)$$

If the volume to capacity ratio ( $X$ ) is larger than 1, then

$$d_a = UD_o + OD = \frac{1}{2}C(1 - g / C) + \frac{T}{2}(X - 1) \quad (6-4)$$

The constraints are composed of three parts: maximum and minimum green; pedestrian setting; and cycle length and NEMA dual ring structure. For more detailed description of bus delay estimation, and constraints elaborated, please refer to our previous research [20]. There may not exist an unambiguous optimal solution that minimizes both the total bus delay and maximum control delay simultaneously. Hence, a set of Pareto optimal solutions or non-dominated solutions are sought instead. All these solutions form a Pareto frontier. Based on transportation manager's consideration of other information, an optimal timing plan can be selected.

### 6.2.2 Solution Algorithm

Multi-objective approach needs to search for non-dominated Pareto optimal solutions, which is efficient algorithms. In recent years, a number of researches has studied this problem and developed a number of GA-based multi-objective optimization tools. One promising method is the Non-dominated Sorting Genetic Algorithms – NSGA [121]. NSGA keeps the normal crossover and mutation operator, but uses different selection operator compared with a simple genetic algorithm. Specifically, before a selection, the selection operator ranks the population according to the individual's non-domination. Recently, Deb et al. developed an upgraded NSGA with several major innovations, named NSGA-II. The major improvements include a fast crowded

distance estimation procedure, a fast non-dominated sorting approach, and a simple crowded comparison operator [113]. NSGA-II is used in this study to solve the multi-objective adaptive priority control problem. Figure 6.2 is the flowchart of NSGA-II Algorithm. The major process is as follows.

Step 1: After specifying the algorithm parameters, such as length of gene, size of population, probability of mutation and crossover, and maximum generation number, the algorithm starts by building a population of individuals based on all the signal phase schemes of the intersections in the corridor.

Step 2: Each individual is evaluated, ranked, and sorted according to the dominance rule.

Step 3: It applies the crossover and mutation operations to create a new population of offs-springs.

Step 4: The parent population and children population are combine to a new population for forming Pareto fronts.

Step 5: The crowding distance is added to each individual, so the algorithm can ensure the diversity of the front. After implementing the fast non-dominated sorting approach, the algorithm obtains the non-dominated fronts of the population.

Step 6: If is meets the criteria of maximum generation, the process stops and save the final children population.

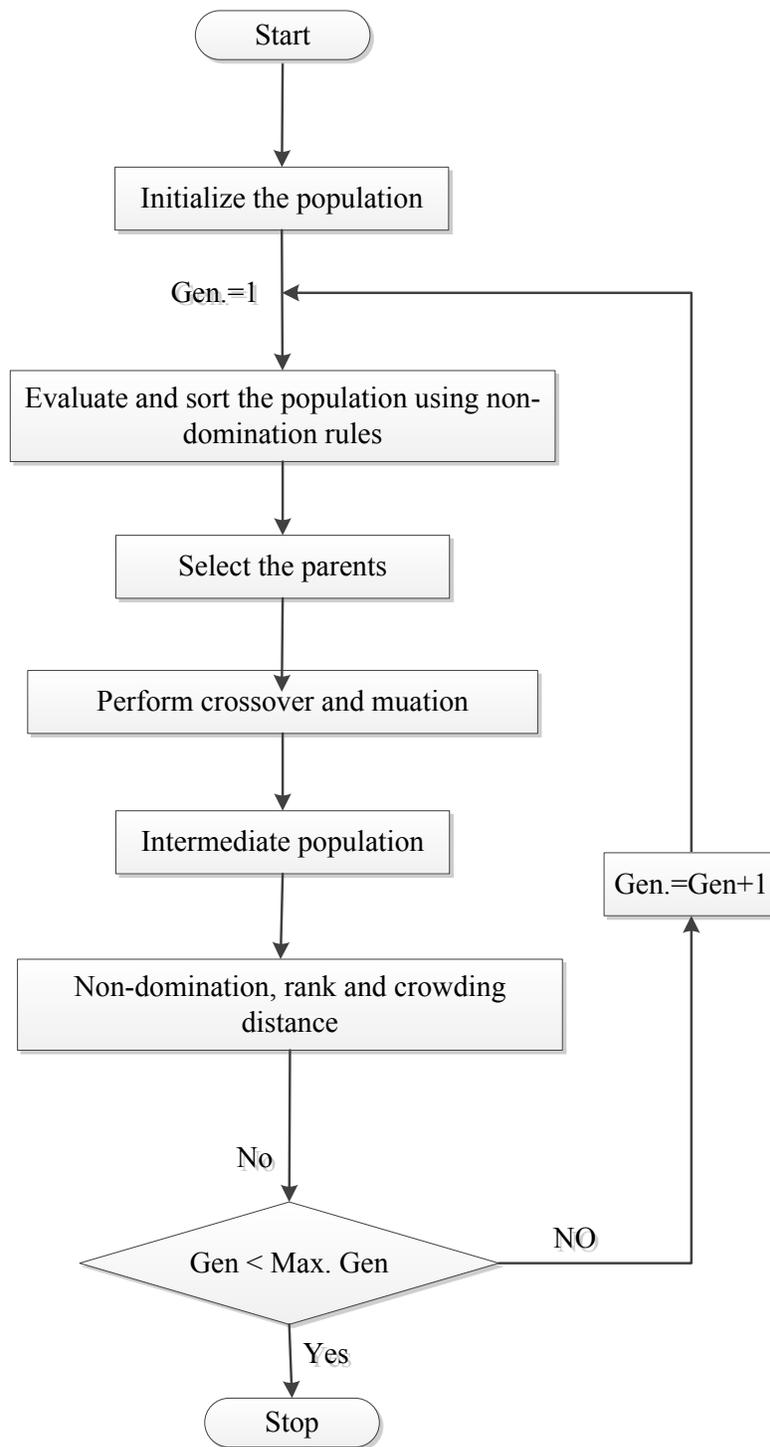


Figure 6.2 Flowchart of NSGA-II Algorithm

### **6.2.3 Solution Selection**

As shown in Figure 6.1, after a set of trade-off solutions are worked out, decision maker will use high-level information to evaluate obtained trade-off solutions and select one solution. Regarding the adaptive priority control problem, advanced technologies have been developed to provide real-time information about the traffic flow conditions and operation information of transit vehicles. For example, traffic states data (e.g. volume and speed) can be collected in real time by inductive loop detector, video, and magnetic sensors placed near the intersection. Automated Vehicle Location (AVL) technologies can track transit vehicles and send the location information continuously. Automatic Passenger Counters (APC) can detect boarding and alighting passengers at transit stops, which can be used to estimate the passenger occupancy of each transit vehicle. To find the final solution of the adaptive priority control problem, this study uses the priority list in Figure 6.3 as an example to explain the decision process.

Step 1: The first priority is to check whether the maximum control delay on one approach exceeds a user defined value. It reflects the overall implications of a control scheme on traffic flow that overflow are controlled to avoid the queue length exceeding a link's storage capacity, which indicates a possible spillback into the preceding intersection. The unsatisfied solutions will be deleted. If all the solutions do not meet this measure, the priority will be inhibited.

Step 2: check schedule adherence (early, on time, or late) to improve public transit schedule consistency. If the bus is behind the schedule, the solution that

minimizes the bus delay will be chosen; if the bus is on time or ahead of the schedule, it will go to the next step.

Step 3: considering the occupancy, select the solution that maximizes personal throughput for transit and private vehicles.

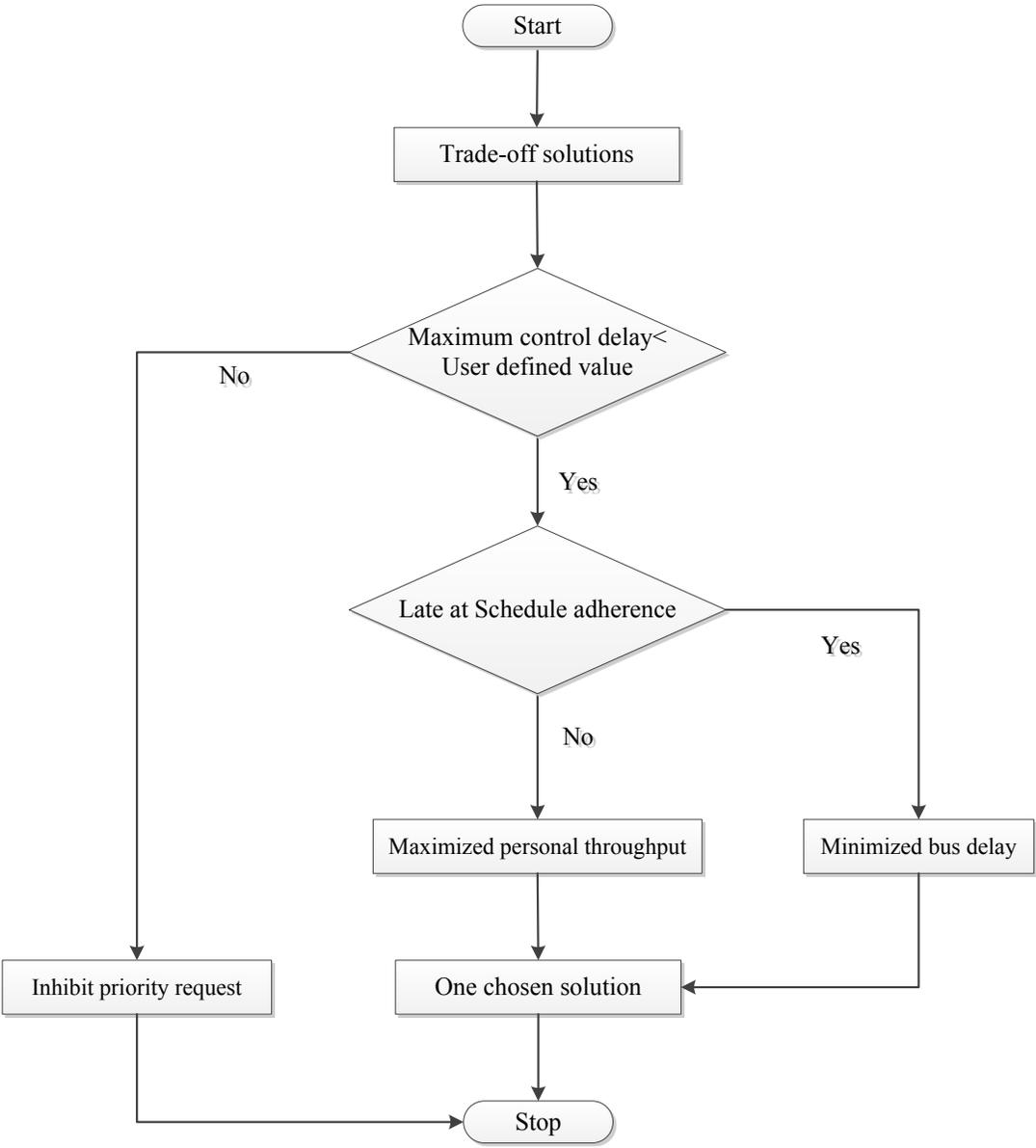


Figure 6.3 One Example of User Prioritized Rules

### 6.3 Simulation Evaluation

The NSGA-II was run to optimize the two objectives: control delay and bus delay.

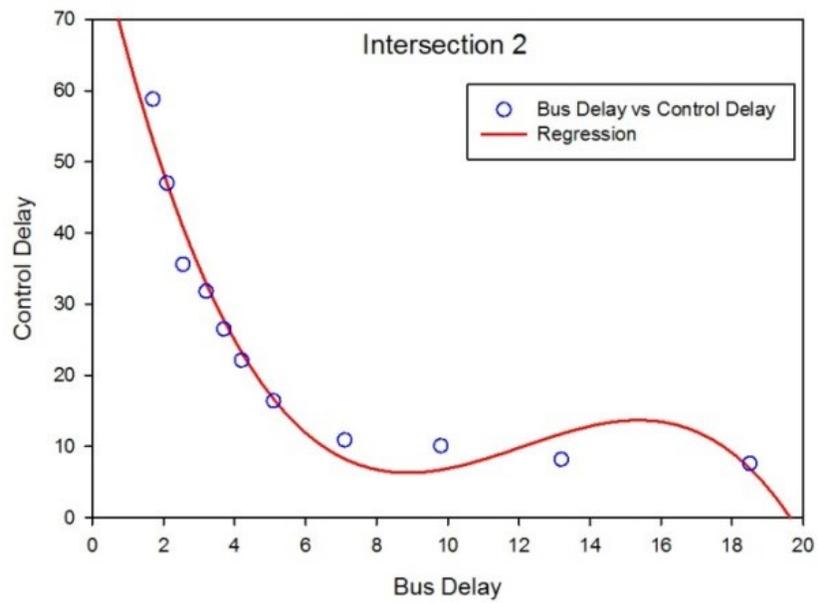
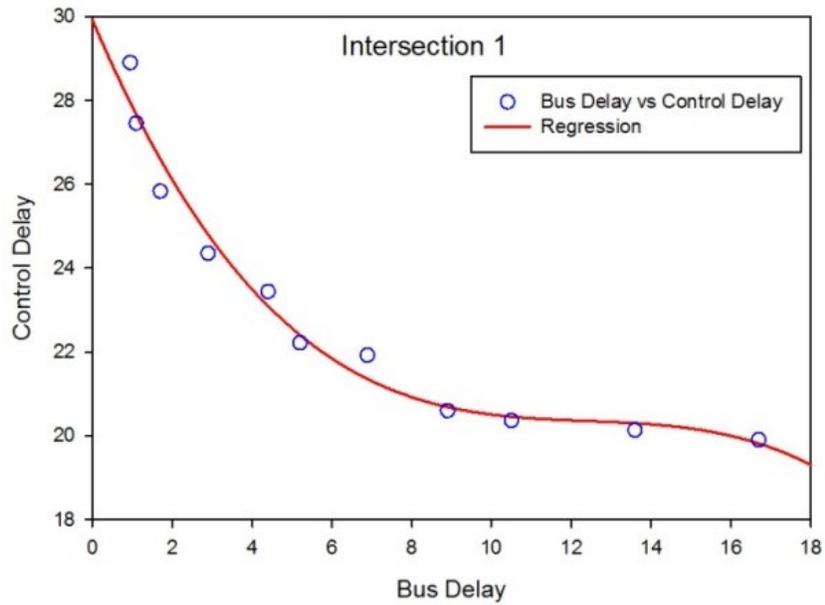
Table 6.2 summarizes the parameters relative to the NSGA-II procedure.

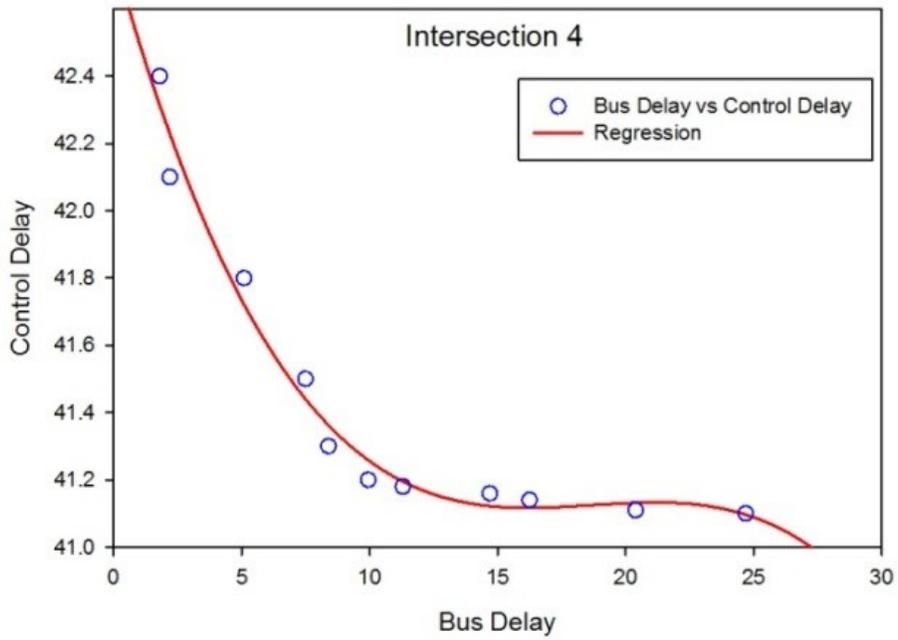
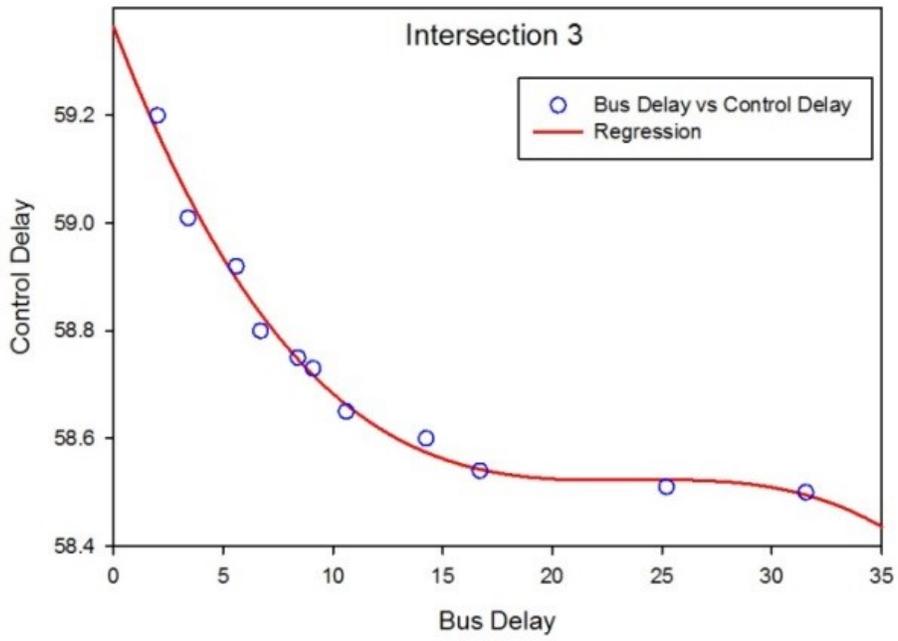
Table 6.2 NSGA-II Parameter Used in Simulation Tests

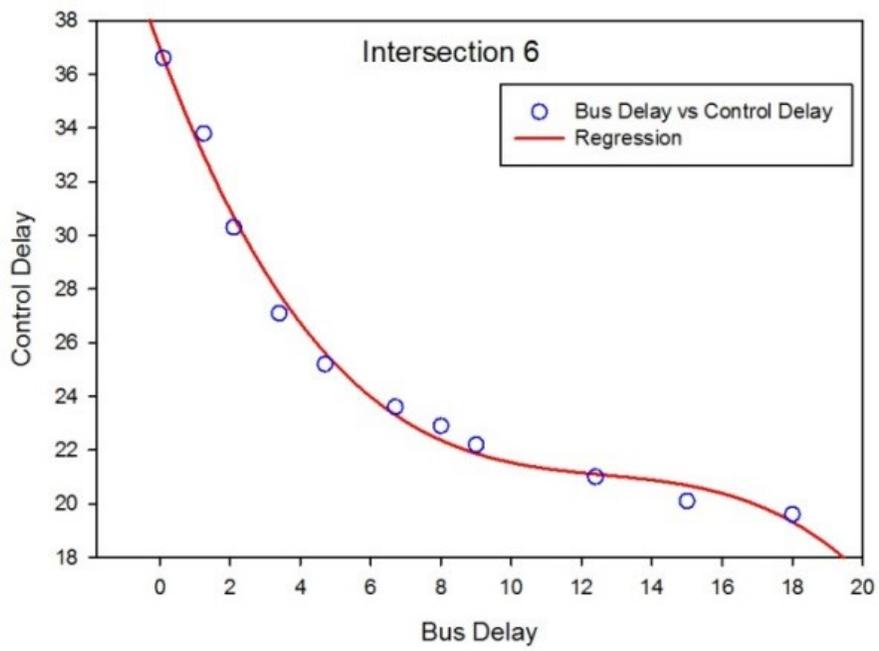
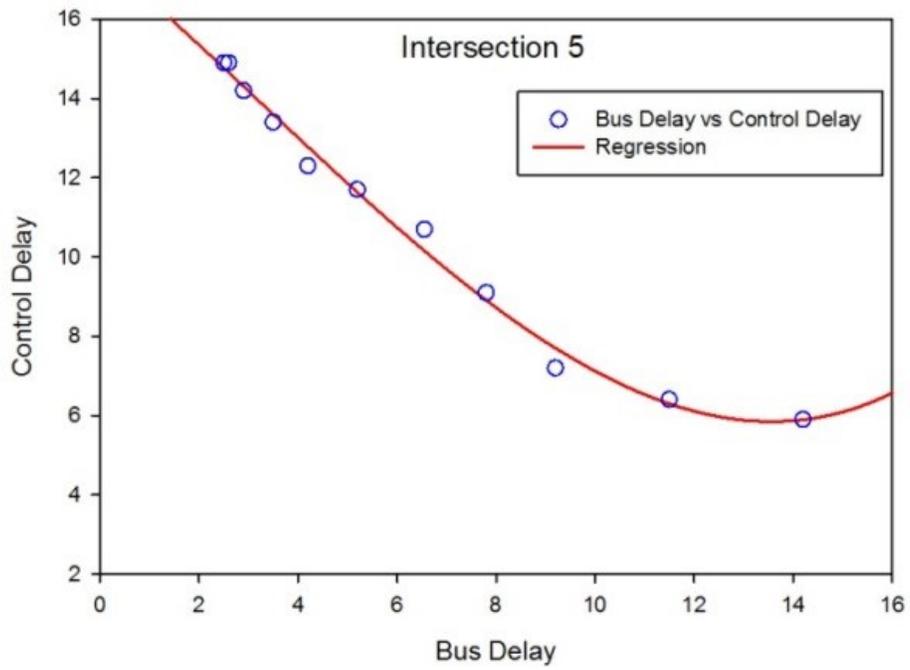
<b>Parameter</b>	<b>Value</b>	<b>Parameter</b>	<b>Value</b>
Population Size	60	No. of Functions	2
Chromosome Length	50	No. of Generation	50
Selection Strategy	Tournament Selection	Cross-over Probability	0.95
Cross-over Probability	0.90	Mutation Probability	0.10
		X-over on binary string	Single point X-over

It is observed that a clear frame of actual Pareto Frontiers is located in the generation 12. As the generation number grows, more Pareto Frontiers are discovered. Figure 6.7 show the relationship between the two objective values at generation 20. The figure clearly shows the trade-offs between control delay and bus delay. Firstly, an obvious conclusion is that bus delay conflicts with the maximum control delay. When the bus delay decreases because of a higher weighting given to the bus, the control delay increases at increasing rates. Second, the trade-off degree is obviously different for the eight intersections, which may depend on the saturation rate of each intersection. In order to evaluate the trade-off between two objectives, a set of well-fitted third degree polynomial regression functions are presented in Table 6.3. It can be observed that the minimum value of delay belongs to the location

where the marginal bus delay and traffic control delay are equal. The best values are different for each intersection. Therefore, the different values should be determined and used at each specific intersection.







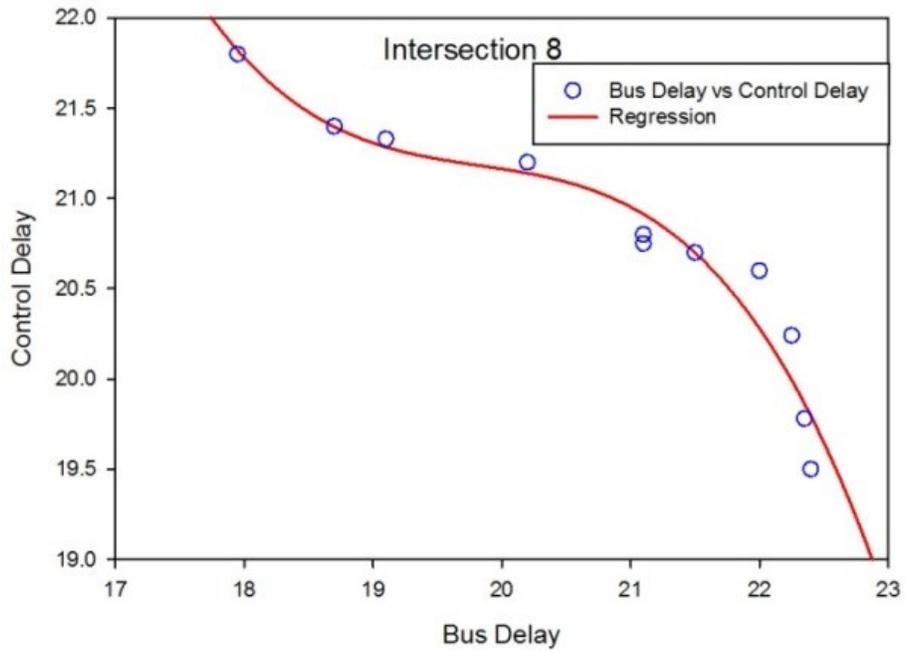
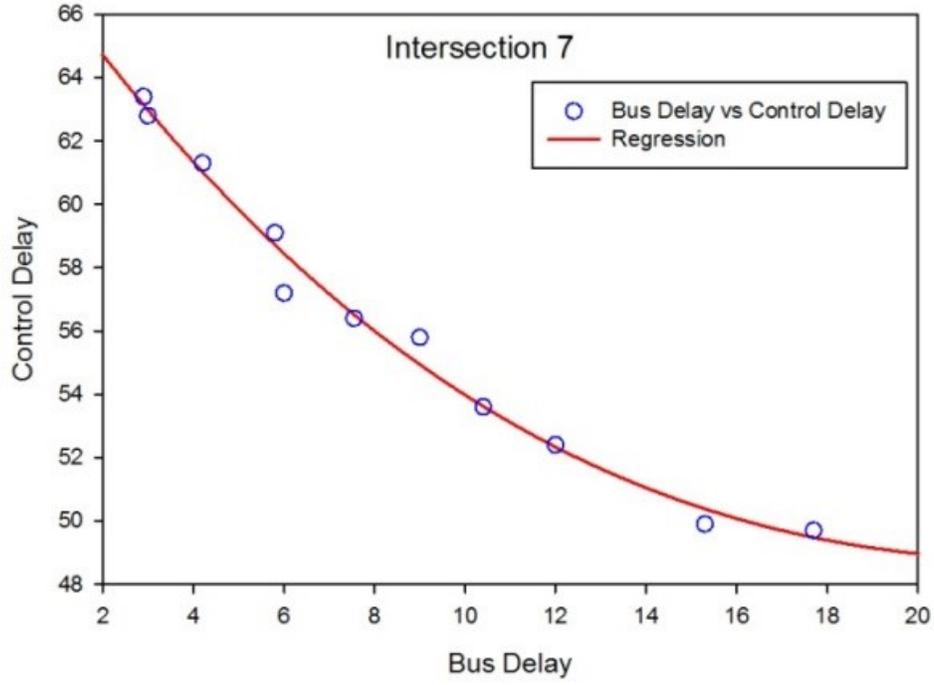


Figure 6.4 Pareto Frontier of Generation 20 from NSGA-II

Table 6.3 Trade-off between Two Objectives

Case	Coefficients				
	b0	b1	b2	b3	r <sup>2</sup>
Intersection 1	29.920	2.241	0.178	0.005	0.977
Intersection 2	85.003	-21.967	1.953	-0.054	0.967
Intersection 3	59.365	-0.107	0.005	0.000	0.989
Intersection 4	42.759	-0.272	0.015	0.000	0.979
Intersection 5	17.585	-1.061	-0.035	0.004	0.992
Intersection 6	36.976	-3.521	0.268	-0.007	0.992
Intersection 7	68.621	-2.082	0.069	-0.001	0.988
Intersection 8	537.763	-7.933	3.924	-0.066	0.938

To analyze the validity of multi-objective optimization method, three control scenarios are considered in this study: (1) baseline: actuated control at signalized intersections without any TSP control strategy; (2) weighted TSP: actuated control at signalized intersections, with the adaptive TSP strategy using weighted combination method; (3) multi-objective TSP: actuated control at signalized intersections, with the adaptive TSP strategy using proposed multi-objective optimization method.

The evaluation and comparison are divided into two categories: (1) the corridor level, which focuses on the total bus travel time and average bus delay along the 7.4 kilometer bus corridor, and (2) individual intersections, including traffic control delay, bus delay, etc. Five Measures of Effectiveness (MOEs) were considered and analyzed: (1) total bus travel time along the corridor; (2) bus delay at

each intersection; (3) control delay at each intersection; (4) schedule adherence along the corridor; (5) personal throughput at each selection. These five MOEs determine priority control performance and benefit.

Table 6.4 shows the statistic tests of the total bus travel time along the corridor. Only the buses driving through the entire corridor are counted as the object. The results are the average value of the multiple runs. In the t-test, one assumption is that the sample of the results follows the normal distribution at a 95% confidence level. Comparing the non-TSP scenario to the TSP scenario, both weighted and multi-objective TSP bring significant bus travel time savings. The mean value of the total travel time shows weighted TSP will save about 60-80 seconds compared to the baseline, and the multi-objective TSP can save 40-95 seconds along the whole corridor compared to the baseline. Multi-objective TSP shows better travel time savings on southbound, as it gives much more priority the buses southbound to improve schedule adherence.

Table 6.4 Total Bus Travel Times along the Corridor

Control Type	Southbound		Northbound	
	Average total travel time (s)	Time saving	Average total travel time (s)	Time saving
Baseline	1087.6	N/A	1081.5	N/A
Weighted	1008.3	79.3	1021.4	60.1
Multi-objective	992.6	95.0	1038.7	42.8
<i>t value</i>	4.96		3.08	
<i>t critical value (two tail)</i>	2.13		2.13	
<i>Confidence Level</i>	95%		95%	
Significant improvement?	Yes		Yes	

In Table 6.5, it can be found that in most of intersections, the multi-objective strategy saves more bus delay than weighted method. Second, it also can be concluded that there are similar reductions on bus delay caused by both two strategies at most of intersections. That means, in terms of bus delay, two methods provide almost the same benefit to the bus. As a trade-off between the bus priority and general traffic delay, two methods consider both granting TSP priority and minimizing general traffic impact. The benefits of multi-objective optimization method are: (1) the personal delay is significantly decreased compared with weighted method; (2) transit serviceability and schedule reliability are improved,

which can result in increased ridership and customer satisfaction; (3) there is no control delay exceeding 50 seconds.

Table 6.5 MOEs at Individual Intersections

<b>Intersection</b>	<b>1</b>			<b>2</b>			<b>3</b>		
MOE	Personal delay	Bus delay	Control delay	Personal delay	Bus delay	Control delay	Personal delay	Bus delay	Control delay
Baseline	18.5	11.0	24.0	15.2	7.3	34.2	20.3	15.8	58.7
Weighted	16.5	8.0	23.4	8.9	4.7	33.0	12.8	9.3	52.5
Difference	2.0	3.0	0.6	6.3	2.6	1.3	7.5	6.5	6.2
Multi-objective	15.2	8.5	25.2	6.5	3.8	33.2	13.2	9.6	51.3
Saving	3.3	2.5	-1.2	8.7	3.5	1.0	7.1	6.2	7.4
<b>Intersection</b>	<b>4</b>			<b>5</b>			<b>6</b>		
MOE	Personal delay	Bus delay	Control delay	Personal delay	Bus delay	Control delay	Personal delay	Bus delay	Control delay
Baseline	19.8	10.4	43.1	6.9	6.2	7.9	29.2	25.4	33.6
Weighted	15.2	7.2	41.5	6.8	6.2	7.2	28.3	20.4	27.5
Difference	4.6	3.2	1.6	0.1	0.0	0.7	0.9	5.0	6.1
Multi-objective	12.3	5.6	42.3	6.2	5.8	7.0	26.1	18.6	26.9
Saving	7.5	4.8	0.8	0.7	0.4	0.9	3.1	5.0	6.1
<b>Intersection</b>	<b>7</b>			<b>8</b>			Corridor		
MOE	Personal delay	Bus delay	Control delay	Personal delay	Bus delay	Control delay	Schedule adherence		
Baseline	32.3	6.4	60.5	24.3	25.0	20.2	83%		
Weighted	28.6	5.3	52.7	19.6	18.1	22.0	83%		
Difference	3.7	1.1	7.8	4.7	6.9	-1.8	0%		
Multi-objective	26.8	5.6	47.6	16.3	15.3	23.1	100%		
Saving	5.5	0.8	12.9	8.0	9.7	-2.9	17%		

## 6.4 Summary

The performance of adaptive priority control depends on three factors: delay estimation, weights determination and optimization formulation. This study used multi-objective optimization method to investigate the priority control performance from the aspects of the weights determination and optimization formulation. It can be concluded from the results that the multi-objective genetic algorithm had potential use in intersection adaptive signal timing optimization. It demonstrated that NSGA-II is efficient to solve multiobjective signal timing design problems under real traffic arrival patterns. Further, the proposed Pareto-frontier regression functions provided an insight into the trade-off among multiple signal optimization objectives.

A set of Pareto optimal signal timing plans are generated that form an efficient frontier. The frontier exhibits an obvious tradeoff between maximum control delay and total bus delay, providing a foundation for the decision making with high-level information. This study proposed prioritized rules for the multi-objective priority control problem. The results showed that multi-objective optimization method can gain better comprehensive traffic benefits than weighted method. The findings provides traffic manager an easy way to select the most appropriate adaptive priority control solutions for particular situations that best serve the needs of transit vehicles and general traffic.

## **Chapter 7 Conclusions and Future Work**

### **7.1 Conclusions**

If the traffic state comes to the realm of congested condition, traffic intersections are not isolated and the traffic states of roads will interact with each other. Hence, it is necessary to understand the behavior of arterial traffic and to investigate coordinated signal control strategies. Model-based adaptive signal control is a promising control methodology that can meet the needs for controlling and coordinating congested arterial traffic. In the thesis, several methods were proposed to address the problems arising when model-based adaptive control methodology is used for signal timing optimization for congested mixed arterial traffic. The main methods considered in the thesis can be summarized as follows.

#### (1) Hierarchical control structure

This study utilized a hierarchical control structure to divide the signal control problem of a large traffic system into three different control layers. Control problems with different details were addressed in different layers: the highest layer optimized the cycle length on the basis of flow capacities and volume levels; the mid layer continuously calculated optimal split with an embedded enhanced SFM and using the rolling horizon scheme for proactive control; the lowest layer adjusted the offsets from the network level by introducing the boundary offsets and considering the spillback offset and starvation offset; there was an extra multi-modal priority control layer to provide priority to different travel modals in the mixed arterial traffic.

## (2) MPC Controller

MPC controller was built to address multiple control problems for arterial traffic corridor. It combined an enhanced SFM, mathematical optimization and rolling-horizon scheme to capture queue interactions among neighboring lane groups in a link and multiple signal phase operation.

## (3) Multi-objective Methodology

This study adopted preemptive goal programming approach to tuning offset in real time for congested arterial corridor, which considered maximum offset and minimum offsets to avoid spillback and starvation. Under mixed arterial traffic, different travel modes competed for the same road space. Priority control was established to favor one mode over another. A new multi-objective optimization problem was formulated to find trade-off solutions between control delay of general traffic and bus delay.

## (4) Solution algorithms based Modified GA

GA is adopted to solve the problem. Many previous studies have also shown the effectiveness of GA when solving signal optimization problems. Modified GA-based signal optimization programs were developed and evaluated.

## (5) SILS based Adaptive Control Implementation

In order to implement and evaluate the adaptive control algorithms, a simulation platform was developed, containing SILS environment, ASC/3 interface, and control system. The adaptive control was implemented in the virtual ASC/3 controller by adding input–output functions over the NTCIP and Transmission Control

Protocol/Internet Protocol (TCP/IP). The signal timings were modified through the proposed optimization models.

#### (6) Simulation Evaluation

Based on the developed simulation platform, the evaluations of the proposed adaptive traffic signal control strategies were conducted on two case studies in the City of Edmonton, Alberta. Each of the scenarios was simulated multiple times and results were tested for statistical significance. Our findings showed that proposed models outperformed actuated signal timings in increasing throughput, decreasing delay, and preventing queue spillback.

### **7.2 Limitations and Recommendations**

Research on adaptive traffic signal control is extremely challenging. There are some limitations of this study.

- Phase sequence is kept constant in the proposed adaptive control strategies and this study did not investigate the phase sequence optimization, but the performance improvement resulting from different sequences cannot be neglected. A reasonable method needs be developed to address phase sequence optimization with considering the cost of disturbing coordination.
- When the arterial corridor is large, corridor-wide cycle length becomes unsuitable. It is necessary to investigate the cycle length optimization under the congested condition.
- Some of the parameters of the traffic models proposed in the thesis, such as turning rates and shockwave speed, were assumed to be constant. Actually,

these parameters will change with time because of route choice decision, weather conditions, seasonal variations, construction events, incidents and others. However, they were not considered sufficiently in this thesis.

- The enhanced SFM model assumed a fixed turning ratio to simulate vehicles merging into different lane groups, but the lane-change behavior may have big impact on the queue dynamic because of stochastic individual drivers.

There are several needs for future research that were identified, including the following:

- With advances in traffic sensing technologies, innovative data sources are available, such as smartphone data and connected vehicle data. These new data source will enhance the capability and accuracy of arterial traffic flow prediction model. Finally, this enhancement can improve the performance of the adaptive control strategies.
- Traffic demand was assumed fixed during the two hour simulation in this study. However, the traffic demand variation was an important nature of arterial traffic network. Therefore, future research can investigate hybrid control models, considering the prediction of the future traffic demand, dynamic traffic assignment, and adaptive signal control,
- All of the experiments were done in simulation environment. It is well known that simulations have certain limitations in representing real-world traffic dynamic under congested condition. Field testing will provide valuable validation of the proposed control strategies.

## References

1. Abu-Lebdeh, G. and R.F. Benekohal, *Design and evaluation of dynamic traffic management strategies for congested conditions*. Transportation Research Part A: Policy and Practice, 2003. 37(2): p. 109-127.
2. Longley, D., *A Control Strategy for A Congested Computer-Controlled Traffic Network*. Transportation Research, 1968. 2(4): p. 391-408.
3. Chaudhary, N.A., et al., *Guidelines for Operating Congested Traffic Singals*, 2010, Texas Department of Transportation.
4. Liu, G., et al., *Development of a Dynamic Control Model for Oversaturated Arterial Corridor*. Procedia - Social and Behavioral Sciences, 2013. 96: p. 2884-2894.
5. Gartner, N.H., J.D. Little, and H. Gabbay, *Optimization of Traffic Signal Settings by Mixed-Integer Linear Programming Part I: The Network Coordination Problem*. Transportation Science, 1975. 9(4): p. 321-343.
6. Gartner, N.H., J.D. Little, and H. Gabbay, *Optimization of Traffic Signal Settings by Mixed-Integer Linear Programming: Part II: The Network Synchronization Problem*. Transportation Science, 1975. 9: p. 344-363.
7. Gartner, N.H., et al., *A Multiband Approach to Arterial Traffic Signal Optimization*. Transportation Research Part B-Methodological, 1991. 25(1): p. 55-74.
8. Smaglik, E., D. Bullock, and T. Urbanik, *Evaluation of Lane-by-Lane Vehicle Detection for Actuated Controllers Serving Multilane Approaches*.

- Transportation Research Record: Journal of the Transportation Research Board, 2005. 1925: p. 123-133.
9. Zheng, X. and L.Y. Chu, *Optimal Parameter Settings for Adaptive Traffic-Actuated Signal Control*. Proceedings of the 11th International IEEE Conference on Intelligent Transportation Systems 2008, New York. 105-110.
  10. NCHRP, *Adaptive Traffic Control Systems: Domestic and Foreign State of Practice*, 2010: US.
  11. Liu, G., et al. *Adaptive Model-Based Offset Optimization for Congested Arterial Network*. in *Transportation Research Board 93rd Annual Meeting*. 2014.
  12. Haj-Salem, H., et al. *Metacor: A Macroscopic Modelling Tool for Urban Corridor*. in *Toward an Intelligent Transport System. Proceedings of the First World Congress on Applications of Transport Telematics and Intelligent Vehicle-Highway Systems, November 30 December 1994, PARIS. Volume 3*. 1994.
  13. Barisone, A., et al. *A macroscopic traffic model for real-time optimization of signalized urban areas*. in *Proceedings of the 41st IEEE Conference on Decision and Control*. 2002.
  14. Head, L., et al., *Modeling Traffic Signal Operations with Precedence Graphs*. Transportation Research Record: Journal of the Transportation Research Board, 2007. 2035: p. 10-18.

15. Lo, H.K. and H.F. Chow, *Adaptive traffic control system: Control strategy, prediction, resolution, and accuracy*. Journal of Advanced Transportation, 2002. 36(3): p. 323-347.
16. Aboudolas, K., M. Papageorgiou, and E. Kosmatopoulos, *Store-and-forward based methods for the signal control problem in large-scale congested urban road networks*. Transportation Research Part C: Emerging Technologies, 2009. 17(2): p. 163-174.
17. He, Q., K.L. Head, and J. Ding, *PAMSCOD: Platoon-based arterial multi-modal signal control with online data*. Transportation Research Part C: Emerging Technologies, 2012. 20(1): p. 164-184.
18. Xie, X.F., et al., *Platoon-Based Self-Scheduling for Real-Time Traffic Signal Control*, in *2011 14th International Ieee Conference on Intelligent Transportation Systems2011*, Ieee: New York. p. 879-884.
19. He, Q., K.L. Head, and J. Ding, *Multi-modal traffic signal control with priority, signal actuation and coordination*. Transportation Research Part C: Emerging Technologies, 2014. 46: p. 65-82.
20. Han, X., et al., *Development and Evaluation of Adaptive Transit Signal Priority Control with Updated Transit Delay Model*. Transportation Research Record: Journal of the Transportation Research Board, 2014. 2438: p. 45-54.
21. Li, M., et al., *Modeling and implementation of adaptive transit signal priority on actuated control systems*. Computer-Aided Civil and Infrastructure Engineering, 2011. 26(4): p. 270-284.

22. Lee, J., A. Shalaby, and B. Abdulhai. *Optimized Strategy for integrated TRAffic and TRAnsit signal Control*. in *4th International Gulf Conference on Roads, November 10, 2008 - November 13, 2008*. 2008. Doha, Qatar: CRC Press.
23. Christofa, E. and A. Skabardonis, *Traffic Signal Optimization with Application of Transit Signal Priority to an Isolated Intersection*. Transportation Research Record: Journal of the Transportation Research Board, 2011. 2259: p. 192-201.
24. Stevanovic, A., P. Martin, and J. Stevanovic, *VisSim-Based Genetic Algorithm Optimization of Signal Timings*. Transportation Research Record: Journal of the Transportation Research Board, 2007. 2035: p. 59-68.
25. Park, B., C. Messer, and T. Urbanik, *Traffic Signal Optimization Program for Oversaturated Conditions: Genetic Algorithm Approach*. Transportation Research Record: Journal of the Transportation Research Board, 1999. 1683: p. 133-142.
26. Zhang, L., Y. Yin, and Y. Lou, *Robust Signal Timing for Arterials Under Day-to-Day Demand Variations*. Transportation Research Record: Journal of the Transportation Research Board, 2010. 2192: p. 156-166.
27. Liu, G., et al., *Optimization of snow plowing cost and time in an urban environment: A case study for the City of Edmonton*. Canadian Journal of Civil Engineering, 2014. 41(7): p. 667-675.

28. Pavlis, Y. and W.W. Recker, *Inconsistencies in the problem of optimal signal control for surface street networks*, in *Itsc 2004: 7th International Ieee Conference on Intelligent Transportation Systems, Proceedings2004*, Ieee: New York. p. 361-366.
29. Gartner, N.H., C.J. Messer, and A.K. Rathi, *Traffic flow theory: A state-of-the-art report2001*: Committe on Traffic Flow Theory and Characteristics (AHB45).
30. Daganzo, C.F., *The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory*. *Transportation Research Part B: Methodological*, 1994. 28(4): p. 269-287.
31. Daganzo, C.F., *The cell transmission model, part II: Network traffic*. *Transportation Research Part B: Methodological*, 1995. 29(2): p. 79-93.
32. Lo, H.K., *A novel traffic signal control formulation*. *Transportation Research Part A: Policy and Practice*, 1999. 33(6).
33. Lo, H.K., *A Cell-Based Traffic Control Formulation: Strategies and Benefits of Dynamic Timing Plans*. *Transportation Science* 2001. 35(2).
34. Lo, H.K., E. Chang, and Y.C. Chan, *Dynamic network traffic control*. *Transportation Research Part a-Policy and Practice*, 2001. 35(8): p. 721-744.
35. Lo, H. and A. Chow, *Control Strategies for Oversaturated Traffic*. *Journal of Transportation Engineering*, 2004. 130(4): p. 466-478.

36. Wei-Hua, L. and W. Chenghong, *An enhanced 0-1 mixed-integer LP formulation for traffic signal control*. Intelligent Transportation Systems, IEEE Transactions on, 2004. 5(4): p. 238-245.
37. Li, Z., *Modeling Arterial Signal Optimization with Enhanced Cell Transmission Formulations*. Journal of Transportation Engineering, 2011. 137(7): p. 445-454.
38. Gazis, D.C., *Optimal Control of a system of Oversaturated Intersections*. Operations Research, 1964. 12(6): p. 815-831.
39. Papageorgiou, M., *An Integrated Control Approach for Traffic Corridors*. Transportation Research Part C-Emerging Technologies, 1995. 3(1): p. 19-30.
40. Aboudolas, K., et al., *A rolling-horizon quadratic-programming approach to the signal control problem in large-scale congested urban road networks*. Transportation Research Part C: Emerging Technologies, 2010. 18(5): p. 680-694.
41. Aboudolas, K., M. Papageorgiou, and E. Kosmatopoulos. *Control and Optimization Methods for Traffic Signal Control in Large-scale Congested Urban Road Networks*. in *American Control Conference, 2007. ACC '07*. 2007.
42. Glomb, A.J., *Dispersion of traffic platoons*, 1989, The University of Arizona: United States -- Arizona. p. 139-139 p.
43. Michalopoulos, P.G. and V. Pisharody, *Platoon Dynamics on Signal Controlled Arterials*. Transportation Science, 1980. 14(4): p. 365-396.

44. Yu, L. *Platoon dispersion and calibration under advanced traffic control strategies*. in *Proceedings of the 1997 Conference on Traffic Congestion and Traffic Safety in the 21st Century, June 8, 1997 - June 11, 1997*. 1997. Chicago, IL, USA: ASCE.
45. Rakha, H. and M. Farzaneh. *Macroscopic modeling of traffic dispersion: Issues and proposed solutions*. in *Transportation Research Board Annual Meeting*. 2005.
46. Farzaneh, M. and H. Rakha, *Procedures for calibrating TRANSYT platoon dispersion model*. *Journal of Transportation Engineering-Asce*, 2006. 132(7): p. 548-554.
47. Hunt, P., et al., *The SCOOT on-line traffic signal optimisation technique*. *Traffic Engineering & Control*, 1982. 23(4).
48. Maher, M., *A comparison of the use of the cell transmission and platoon dispersion models in TRANSYT 13*. *Transportation Planning and Technology*, 2011. 34(1): p. 71-85.
49. Manar, A. and K. Baass, *Traffic Platoon Dispersion Modeling on Arterial Streets*. *Transportation Research Record: Journal of the Transportation Research Board*, 1996. 1566: p. 49-53.
50. Yu, L., *Calibration of Platoon Dispersion Parameters on the Basis of Link Travel Time Statistics*. *Transportation Research Record: Journal of the Transportation Research Board*, 2000. 1727: p. 89-94.

51. Rakha, H. and M. Farzaneh, *Issues and solutions to macroscopic traffic dispersion modeling*. Journal of Transportation Engineering, 2006. 132(7): p. 555-564.
52. S.C. Wong, W.T.W., Jianmin Xu and C.O. Tong, *A Time-dependent TRANSYT Traffic Model for Area Traffic Control*. TRAFFIC AND TRANSPORTATION STUDIES, 2000.
53. Gartner, N.H., M. Kaltenbach, and M. Miyamoto, *Demand-Response Decentralized Urban Traffic Control: Part 2. Network Extensions*, 1983.
54. Gartner, N.H., *demand-response decentralized urban traffic control part I - single intersection politics*, 1982.
55. Farges, J.-L., Henry, J.-J., and Tufal, J. , *The PRODYN real-time traffic algorithm*, in *4th IFAC Symp. Transportation Systems* 1983.
56. Boillot, F., S. Midenet, and J.-C. Pierrelée, *The real-time urban traffic control system CRONOS: Algorithm and experiments*. Transportation Research Part C: Emerging Technologies, 2006. 14(1): p. 18-38.
57. Bielefeldt, C. and F. Busch. *MOTION-a new on-line traffic signal network control system*. in *Road Traffic Monitoring and Control, 1994., Seventh International Conference on*. 1994.
58. Donati, F., et al. *A hierarchical decentralized traffic light control system*. in *The First Realisation: 'Progetto Torino'*. In: *IFAC 9th World Congress of the International Federation of Automatic Control, Budapest*. 1984.

59. Papageorgiou, M., et al., *Review of road traffic control strategies*. Proceedings of the IEEE, 2003. 91(12): p. 2043-2067.
60. Gartner, N.H., F.J. Pooran, and C.M. Andrews. *Implementation of the OPAC adaptive control strategy in a traffic signal network*. in *Intelligent Transportation Systems, 2001. Proceedings. 2001 IEEE*. 2001.
61. Mirchandani, P. and L. Head, *A real-time traffic signal control system: architecture, algorithms, and analysis*. Transportation Research Part C: Emerging Technologies, 2001. 9(6): p. 415-432.
62. Mirchandani, P. and W. Fei-Yue, *RHODES to intelligent transportation systems*. Intelligent Systems, IEEE, 2005. 20(1): p. 10-15.
63. Pohlmann, T. and B. Friedrich. *Online control of signalized networks using the Cell Transmission Model*. in *13th International IEEE Annual Conference on Intelligent Transportation Systems, Madeira Island, Portugal*. 2010.
64. G.C. Dans, D.C.G., *Optimal Control of Oversaturated Store-and Forward Transportation Networks*. Transportation Science, 1976.
65. Diakaki, C., M. Papageorgiou, and T. McLean, *Integrated Traffic-Responsive Urban Corridor Control Strategy in Glasgow, Scotland: Application and Evaluation*. Transportation Research Record: Journal of the Transportation Research Board, 2000. 1727: p. 101-111.
66. Diakaki, C., M. Papageorgiou, and K. Aboudolas, *A multivariable regulator approach to traffic-responsive network-wide signal control*. Control Engineering Practice, 2002. 10(2): p. 183-195.

67. Diakaki, C., et al., *Extensions and New Applications of the Traffic-Responsive Urban Control Strategy: Coordinated Signal Control for Urban Networks*. Transportation Research Record: Journal of the Transportation Research Board, 2003. 1856: p. 202-211.
68. Timmermans, W., P. Van Den Bosch, and J. Klijnhout, *Improved Network Control by Using TRANSYT*. Traffic Engineering & Control, 1979. 20(7).
69. Zhang, Y. and Y. Xie. *Comparison of PASSER, Synchro, and TRANSYT-7F for Arterial Signal Timing based on CORSIM Simulation*. in *Applications of Advanced Technology in Transportation. Proceedings of the Ninth International Conference*. 2006.
70. Bretherton, D., M. Bodger, and N. Baber. *SCOOT - the future in Road Transport Information and Control, 2004. RTIC 2004. 12th IEE International Conference on*. 2004.
71. Wey, W.-M. and R. Jayakrishnan, *Network traffic signal optimization formulation with embedded platoon dispersion simulation*. Transportation Research Record, 1999(1683): p. 150-159.
72. Li, Z., *An integrated control model for freeway interchanges*, 2011, University of Maryland, College Park: United States -- Maryland. p. 152.
73. Gartner, N., J. Little, and H. Gabbay, *Mitrop: a computer program for simultaneous optimisation of offsets, splits and cycle time*. Traffic Engineering & Control, 1976. 17(8/9).

74. Abu-Lebdeh, G. and R. Benekohal, *Development of Traffic Control and Queue Management Procedures for Oversaturated Arterials*. Transportation Research Record: Journal of the Transportation Research Board, 1997. 1603: p. 119-127.
75. Abu-Lebdeh, G. and R. Benekohal, *Signal Coordination and Arterial Capacity in Oversaturated Conditions*. Transportation Research Record: Journal of the Transportation Research Board, 2000. 1727: p. 68-76.
76. Girianna, M. and R. Benekohal, *Dynamic Signal Coordination for Networks with Oversaturated Intersections*. Transportation Research Record: Journal of the Transportation Research Board, 2002. 1811: p. 122-130.
77. Liu, Y. and G.L. Chang, *An arterial signal optimization model for intersections experiencing queue spillback and lane blockage*. Transportation Research Part C-Emerging Technologies, 2011. 19(1): p. 130-144.
78. Lieberman, E. and J. Woo, *SIGOP II: A new computer program for calculating optimal signal timing patterns*. Transportation Research Record, 1976(596).
79. Lieberman, E., J. Chang, and E. Prassas, *Formulation of Real-Time Control Policy for Oversaturated Arterials*. Transportation Research Record: Journal of the Transportation Research Board, 2000. 1727: p. 77-88.
80. für, V.d.F., *A New Offset Optimization Method for Signalized Urban Road Networks*, 2006

81. Cai, C., C.K. Wong, and B.G. Heydecker, *Adaptive traffic signal control using approximate dynamic programming*. Transportation Research Part C: Emerging Technologies, 2009. 17(5): p. 456-474.
82. Lin, F.-B., D. Cooke, and S. Vijayakumar, *Use of Predicted Vehicle Arrival Information for Adaptive Signal Control - An Assessment*. 1987(1112).
83. Stevanovic, A., et al., *Microscopic Modeling of Traffic Signal Operations*. Transportation Research Record: Journal of the Transportation Research Board, 2009. 2128: p. 143-151.
84. Rosetti, R.J. and R. Liu, *Advances in Artificial Transportation Systems and Simulation* 2014: Academic Press.
85. Stevanovic, A., J. Stevanovic, and P. Martin, *Optimizing Signal Timings from the Field: VISGAOST and VISSIM-ASC/3 Software-in-the-Loop Simulation*. Transportation Research Record: Journal of the Transportation Research Board, 2009(2128): p. 114-120.
86. Zhang, X., et al., *Hierarchical fuzzy rule-based system optimized with genetic algorithms for short term traffic congestion prediction*. Transportation Research Part C: Emerging Technologies, 2014. 43, Part 1: p. 127-142.
87. Hu, H. and H.X. Liu, *Arterial offset optimization using archived high-resolution traffic signal data*. Transportation Research Part C: Emerging Technologies, 2013. 37: p. 131-144.

88. Park, B.B., N. Roupail, and J. Sacks, *Assessment of Stochastic Signal Optimization Method Using Microsimulation*. Transportation Research Record: Journal of the Transportation Research Board, 2001. 1748: p. 40-45.
89. Van den Berg, M., et al. *A macroscopic traffic flow model for integrated control of freeway and urban traffic networks*. in *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*. 2003.
90. Yue, L., et al. *A Lane-group Based Macroscopic Model for Signalized Intersections Account for Shared Lanes and Blockages*. in *Intelligent Transportation Systems, 2008. ITSC 2008. 11th International IEEE Conference on*. 2008.
91. Jabari, S.E. and H.X. Liu, *A stochastic model of traffic flow: Gaussian approximation and estimation*. Transportation Research Part B: Methodological, 2013. 47: p. 15-41.
92. Rawlings, J.B., *Tutorial overview of model predictive control*. Control Systems, IEEE, 2000. 20(3): p. 38-52.
93. Liu, Y., *An integrated traffic control system for freeway corridors under non-recurrent congestion*, 2009, University of Maryland, College Park: United States -- Maryland. p. 218.
94. Bonneson, J.A., S.R. Sunkari, and M.P. Pratt, *Traffic Signal Operations Handbook*, 2009, Texas Transportation Institute, Texas A & M University System.

95. Morgan, J.T. and J.D.C. Little, *Synchronizing Traffic Signals for Maximal Bandwidth*. Operations Research, 1964. 12(6): p. 896-912.
96. Little, J.D.C., M.D. Kelson, and N.H. Gartner, *MAXBAND: a program for setting signals on arterials and triangular networks*. Transportation Research Record, 1981. 795: p. 40-46.
97. Chaudhary, N.A. and C.J. Messer, *Passer IV: A Program for Optimizing Signal Timing in Grid Networks (With Discussion and Closure)*, in *72nd Annual Meeting of the Transportation Research Board* 1993: Washington, DC.
98. Yin, Y., M. Li, and A. Skabardonis, *Offline Offset Refiner for Coordinated Actuated Signal Control Systems*. Journal of Transportation Engineering, 2007. 133(7): p. 423-432.
99. Hillier, J.A. and R. Rothery, *The Synchronization of Traffic Signals for Minimum Delay*. Transportation Science, 1967. 1(2): p. 81-94.
100. Robertson, D.I., *RESEARCH ON THE TRANSYT AND SCOOT METHODS OF SIGNAL COORDINATION*. The Journal-Institute of Transportation Engineers, 1986. 56(1): p. 36-40.
101. Day, C., et al., *Evaluation of Arterial Signal Coordination*. Transportation Research Record: Journal of the Transportation Research Board, 2010. 2192: p. 37-49.
102. Day, C., et al., *Reliability, Flexibility, and Environmental Impact of Alternative Objective Functions for Arterial Offset Optimization*.

- Transportation Research Record: Journal of the Transportation Research Board, 2011. 2259: p. 8-22.
103. Hale, D. and K. Courage, *Prediction of Traffic-Actuated Phase Times on Arterial Streets*. Transportation Research Record: Journal of the Transportation Research Board, 2002. 1811: p. 84-91.
  104. Shoup, G. and D. Bullock, *Dynamic Offset Tuning Procedure Using Travel Time Data*. Transportation Research Record: Journal of the Transportation Research Board, 1999. 1683: p. 84-94.
  105. Chang, E., *Guidelines for Actuated Controllers in Coordinated Systems*. Transportation Research Record: Journal of the Transportation Research Board, 1996. 1554: p. 61-73.
  106. Kuzbari, R., *Early Green Start Analysis for Time-of-Day Signal Coordination*. ITE journal, 1996. 66(8): p. 7.
  107. Abbas, M.M. and D. Bullock, *On-line measure of shockwaves for its applications*. Journal of Transportation Engineering, 2003. 129(1): p. 1-6.
  108. Wu, X. and H.X. Liu, *A shockwave profile model for traffic flow on congested urban arterials*. Transportation Research Part B: Methodological, 2011. 45(10): p. 1768-1786.
  109. Lighthill, M.J. and G.B. Whitham, *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, 1955. 229(1178): p. 317-345.

110. Jones, D. and M. Tamiz, *Practical goal programming*. Vol. 141. 2010: Springer.
111. Sherali, H. and A. Soyster, *Preemptive and nonpreemptive multi-objective programming: Relationship and counterexamples*. *Journal of Optimization Theory and Applications*, 1983. 39(2): p. 173-186.
112. Lertworawanich, P., M. Kuwahara, and M. Miska, *A New Multiobjective Signal Optimization for Oversaturated Networks*. *Intelligent Transportation Systems, IEEE Transactions on*, 2011. 12(4): p. 967-976.
113. Deb, K., et al., *A fast and elitist multiobjective genetic algorithm: NSGA-II*. *Evolutionary Computation, IEEE Transactions on*, 2002. 6(2): p. 182-197.
114. Christofa, E., K. Aboudolas, and A. Skabardonis. *Arterial traffic signal optimization: a person-based approach*. in *The 92nd Annual Meeting of the Transportation Research Board, Washington DC, USA*. 2013.
115. Duerr, P., *Dynamic Right-of-Way for Transit Vehicles: Integrated Modeling Approach for Optimizing Signal Control on Mixed Traffic Arterials*. *Transportation Research Record: Journal of the Transportation Research Board*, 2000. 1731: p. 31-39.
116. Stevanovic, A., et al. *Traffic control optimization for multi-modal operations in a large-scale urban network*. in *Integrated and Sustainable Transportation System (FISTS), 2011 IEEE Forum on*. 2011.
117. Ma, W., Y. Liu, and X. Yang, *A Dynamic Programming Approach for Optimal Signal Priority Control Upon Multiple High-Frequency Bus*

- Requests*. Journal of Intelligent Transportation Systems, 2012. 17(4): p. 282-293.
118. Amer, S., L. Jinwoo, and A. Baher, *Optimized strategy for integrated traffic and transit signal control*, in *Efficient Transportation and Pavement Systems* 2008, CRC Press.
119. Medina, J., E. Lo, and R. Benekohal, *Multiattribute Decision-Making Methods for Optimal Selection of Traffic Signal Control Parameters in Multimodal Analysis*. Transportation Research Record: Journal of the Transportation Research Board, 2014. 2438: p. 64-71.
120. Board, T.R., *Highway Capacity Manual*, 2010, Transportation Research Board.
121. Srinivas, N. and K. Deb, *Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms*. Evolutionary Computation, 1994. 2(3): p. 221-248.