



Minnesota
Department of
Transportation

Improving Traffic Signal Operations for Integrated Corridor Management

**RESEARCH
SERVICES**

Office of
Policy Analysis,
Research &
Innovation

Henry X. Liu, Principal Investigator
Department of Civil Engineering
University of Minnesota

July 2013

Research Project
Final Report 2013-17

Your Destination... Our Priority



To request this document in an alternative format, call Bruce Lattu at 651-366-4718 or 1-800-657-3774 (Greater Minnesota); 711 or 1-800-627-3529 (Minnesota Relay). You may also send an e-mail to bruce.lattu@state.mn.us. (Please request at least one week in advance).

Technical Report Documentation Page

1. Report No. MN/RC 2013-17	2.	3. Recipients Accession No.	
4. Title and Subtitle Improving Traffic Signal Operations for Integrated Corridor Management		5. Report Date July 2013	
		6.	
7. Author(s) Henry X. Liu and Heng Hu		8. Performing Organization Report No.	
9. Performing Organization Name and Address Department of Civil Engineering University of Minnesota 500 Pillsbury Dr. S.E. Minneapolis, MN 55455-0220		10. Project/Task/Work Unit No. CTS Project #2011039	
		11. Contract (C) or Grant (G) No. (C) 89261, (WO) 202	
12. Sponsoring Organization Name and Address Minnesota Department of Transportation Research Services 395 John Ireland Boulevard, Mail Stop 330 St. Paul, MN 55155		13. Type of Report and Period Covered Final Report	
		14. Sponsoring Agency Code	
15. Supplementary Notes http://www.lrrb.org/pdf/201317.pdf			
16. Abstract (Limit: 250 words) The Integrated Corridor Management (ICM) approach has drawn increasingly more attention in recent years because it is believed to be a promising tool to mitigate urban traffic congestion. In this project, a maximum flow based control model was first developed to handle oversaturated traffic conditions at signalized arterials. Based on the arterial control model, an integrated control model was proposed to manage network congestion. Through diversion control, the model aims to fully utilize the available capacity along parallel routes. The impact of the diversion traffic is considered, especially for signalized arterials, so that traffic congestion on the diversion route can be reduced or eliminated by proper adjustment of signal timings. This model does not rely on time-dependent traffic demand as model inputs and it is ready to be implemented at typical parallel traffic corridors where the standard detection system is available. The performance of the proposed model was tested using microscopic traffic simulation in the I-394 and TH 55 corridor in Minneapolis, Minnesota. The results indicate that the proposed model can significantly reduce network congestion.			
17. Document Analysis/Descriptors Integrated corridor management, Control, Maximum flow problem, SMART-SIGNAL, Signal operation		18. Availability Statement No restrictions. Document available from: National Technical Information Services, Alexandria, VA 22312	
19. Security Class (this report) Unclassified	20. Security Class (this page) Unclassified	21. No. of Pages 66	22. Price

Improving Traffic Signal Operations for Integrated Corridor Management

Final Report

Prepared by

Henry X. Liu
Heng Hu

Department of Civil Engineering
University of Minnesota

July 2013

Published by

Minnesota Department of Transportation
Research Services
395 John Ireland Boulevard
Mail Stop 330
St. Paul, MN 55155

This report documents the results of research conducted by the authors and does not necessarily represent the views or policies of the Minnesota Department of Transportation or the University of Minnesota. This report does not contain a standard or specified technique.

The authors, the Minnesota Department of Transportation, and the University of Minnesota do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to this report.

Acknowledgements

This work was supported by Minnesota Department of Transportation. The authors would like to thank Steve Misgen, Timothy Bangsund and Curt Krohn for their assistance in the field deployment of the SMART-Signal system on TH 55. We also want to thank Dr. Darcy Bullock and Howell Li of Purdue University, Jim Sturdevant of Indiana Department of Transportation, and Marc Miranda of Econolite, for their assistance on the data logger function of the Econolite ASC/3 controller.

Dr. Henry Liu and the University of Minnesota have equity and royalty interests in SMART Signal Technologies, Inc., a Minnesota-based private company which could commercially benefit from the results of this research. These relationships have been reviewed and managed by the University of Minnesota in accordance with its Conflict of Interests policies.

Table of Contents

Chapter 1. Introduction.....	1
1.1. Project Motivation.....	1
1.2. Project Objectives	1
1.3. Project Overview.....	1
1.4. Report Organization	2
Chapter 2. Review on the Theory and Practice of the Integrated Corridor Management	3
2.1 Background	3
2.1.1 Definitions.....	3
2.1.2 Fundamentals	4
2.2 Model Development.....	5
2.2.1 Focus on information provision and travelers' response	5
2.2.2 Focus on traffic evolution and interaction	6
2.3 Operations	7
2.3.1 ITS Technologies	7
2.3.2 Practices and Pioneer Sites	8
2.4 Challenges and Opportunities	9
2.4.1 On the Modeling Side	9
2.4.2 On the Operation Side.....	9
2.5 Summary	9
Chapter 3. Managing Oversaturated Signalized Arterials: A Maximum Flow Based Approach.	11
3.1 OSI-based Mitigation Strategies	11
3.2 Problem Definition and Assumptions	14
3.3 The Maximum-Flow Program (MFP).....	15
3.3.1 Control Variables	15
3.3.2 Constraint Analysis.....	17
3.3.3 Maximum Flow Based Control Model	19
3.4 Solution Method – A Forward-Backward Procedure (FBP).....	23
3.4.1 FBP for One Oversaturated Route	23
3.4.2 FBP for a Network with Two Intersecting Oversaturated Routes	25
3.4.3 The Optimality Analysis of the Solution of the FBP	26

3.5	Simulation Test	29
Chapter 4.	Develop Integrated Control Strategies.....	35
4.1	Problem Statement	35
4.2	Model Formulation.....	35
4.2.1	Performance Estimation.....	35
4.2.2	Diversion Control.....	38
4.3	Simulation Test	40
Chapter 5.	Installation of the SMART Signal System on TH 55	47
5.1	The New Data Collection Software	47
5.2	Implementation on TH 55	50
Chapter 6.	Concluding Remarks.....	52
References	53

List of Tables

Table 3.1. Traffic flow conditions during the simulation period.....	30
Table 3.2. Control strategy comparison.....	30
Table 3.3. Network performance comparison.....	31
Table 3.4. Two-hour throughputs comparison.....	32
Table 4.1. Network performance comparison.....	44
Table 4.2. Network performance comparison with demand variations.....	46
Table 5.1 Binary data format in a data chunk.....	49
Table 5.2 Econolite ASC/3 event code interpretation.....	50

List of Figures

Figure 3.1 Green extension for Scenario 1.....	12
Figure 3.2 Red extension for Scenario 2.....	13
Figure 3.3 Red reduction (at downstream intersection) for Scenario 3.....	14
Figure 3.4 Intersections along an oversaturated route.....	15
Figure 3.5 Red time changes & green time changes.....	16
Figure 3.6 Signal timing changes ($\Delta r_{n,i}(t) < 0$, $\Delta g_{n,i}(t) > 0$).....	16
Figure 3.7 An example of applying strategies 2 & 3 to eliminate spillover.....	17
Figure 3.8 Maximum flow network for one oversaturated route.....	21
Figure 3.9 Two intersecting oversaturated routes.....	21
Figure 3.10 Maximum flow network for two intersecting oversaturated routes.....	23
Figure 3.11 Vissim simulation network (a) normal flow condition (b) increased flow condition.....	30
Figure 3.12 Comparison of network performance.....	31
Figure 3.13 Comparison of southbound throughputs.....	32
Figure 3.14 Comparison of side streets' maximum queue length in each cycle.....	33
Figure 4.1 Problem statement.....	35
Figure 4.2 Diversion control from freeway to arterial.....	39
Figure 4.3 Case study site: the TH 55/I-394 corridor, Minneapolis, MN.....	40
Figure 4.4 Flow-density diagram from three detectors at segment 4.....	41
Figure 4.5 VISSIM network of the TH 55/I-394 corridor.....	41
Figure 4.6 Demand profiles for the simulation period.....	42
Figure 4.7 Travel times of general route and diverting route under different scenarios.....	43
Figure 4.8 Travel time and diversion rate.....	43
Figure 5.1 Configuration file.....	47
Figure 5.2 Flow chart for data retrieving software.....	48
Figure 5.3 Sample raw data (<i>The binary data can't be displayed directly based on ANSI coding</i>).....	49
Figure 5.4 Data trunk sample in decimal format.....	49
Figure 5.5 Intersection level of service along TH 55.....	51

Executive Summary

Integrated Corridor Management (ICM), considered in the context of Advanced Traffic Management Systems (ATMS), offers a broad while complex operation concept for improving travel efficiency, system reliability, and traffic safety for transportation corridors. The ICM approach has garnered increasingly more attention in recent years because it is believed to be a promising tool for managing urban traffic congestion.

In this project, a new ICM control system is developed to reduce network congestion by utilizing available capacity of parallel routes. Compared with previous control models, the proposed model specifically considers the impact of the diverting traffic to diversion route, especially for signalized arterials. A new maximum flow-based signal control model is developed to manage arterial congestion based on the *oversaturation severity index (OSI)*, so the potential congestion caused by diverting traffic can be reduced or eliminated by proper adjustment of signal timings. The proposed model does not have the requirement of time-dependent traffic demand information as model input. It is ready to be implemented at typical parallel traffic corridors where a standard detection system is available. Because of the very low computation burden, the model is suitable for on-line applications.

To test the performance of the proposed ICM control model, a microscopic traffic simulation was built for this project based on the I-394 and TH 55 corridor in Minneapolis. The results indicate that the proposed model significantly reduces network congestion. In our test cases, network performance measures, such as average delay per vehicle, average number of stops per vehicle, and average vehicle speed, are improved greatly using the proposed model. The ICM control system developed in this project has a very promising future for real field implementation and we look forward to testing the field performance of the proposed approach in future projects.

A new version of the SMART-Signal data collection software was also developed as part of this project, which can directly retrieve event-based traffic data from the Econolite ASC/3 controllers without additional hardware instrumentation. The software was implemented in the 10 intersections along TH 55 in Minneapolis.

Chapter 1. Introduction

1.1. Project Motivation

Due to the increasing traffic demand but limited facility capacities, traffic congestion has become an increasingly severe problem for metropolitan areas not only in the United States, but also around the world. Determining how to efficiently and effectively manage traffic during peak hours or non-recurrent congestion periods appears to be a challenging task for researchers and practitioners. The Integrated Corridor Management (ICM) approach has drawn more and more attention in recent years because it is believed to be a promising tool to manage urban traffic congestion.

The previous development on the SMART-Signal (Systematic Monitoring of Arterial Road Traffic and Signals) system makes it possible to continuously monitor the traffic situation of signalized arterials in a quantitative way. In particular, the system can simultaneously collect event-based high resolution traffic data from signalized intersections and generate performance measures in real time. At single intersection level, the queue length profile on a cycle-to-cycle basis is monitored by the system, and other performance measures, e.g. intersection delay, Oversaturation Severity Indices (OSI) and Level of Service (LOS), can be further calculated. At arterial level, travel time, speed, and average number of stops can be generated and reported in real time. This project will be built upon these research efforts and apply the developed performance measures to traffic signal operations from an integrated system perspective.

1.2. Project Objectives

This project aims to develop an incident-responsive signal control system which can diagnose problems, measure performance and suggest optimized signal control strategies in real time. In particular, this research will utilize the performances measured by the SMART-Signal system to diagnose incidents on signalized arterials and propose new control strategies to mitigate traffic congestion. Also, when an accident occurs on freeway and create traffic congestion, the proposed system will help identify and predict traffic conditions on the arterial, so that proper traffic management on freeway can be deployed to mitigate or postpone freeway traffic congestion.

1.3. Project Overview

The work planned in this project can be described in the following aspects, including:

- a) Literature review
- b) Development of control strategies for oversaturated arterials
- c) Development of integrated control strategy
- d) Simulation test
- e) Field implementation of the SMART-Signal system on TH 55

Literature review

A thorough literature review on the state-of-the-art integrated corridor management strategies will be conducted. It aims to explore the existing ICM schemes and their field evaluations, and summarize the lessons learned from the previous research and implementation efforts.

Development of control strategies for oversaturated arterials

In order to alleviate the detrimental impacts caused by oversaturation at signalized arterials, a new control model will be developed in this task to adjust corresponding signal timings. The focus is to mitigate two types of detrimental effects, signal phase failure with residual queue and downstream queue spillover.

Development of integrated control strategy

An integrated control model will be developed to handle incidents occurring in adjacent freeway corridors. The model will collect incidents and traffic information on freeways, predict traffic demand diverted from freeway to arterials and adjust the signal timings on arterials to minimize delay caused by detour traffic.

Simulation test

A microscopic simulation model for the TH 55 and I-394 corridors will be developed in VISSIM to test the effectiveness of the signal control model proposed in this project.

Field implementation of the SMART-Signal system on TH 55

In this task, the SMART-Signal system will be installed in the 10 intersections along TH 55, Minneapolis, MN. A new version of the SMART-Signal data collection software will be developed to retrieve event-based traffic data from the Econolite ASC/3 controllers without additional hardware instrument.

1.4. Report Organization

The rest of this report is organized as following: The review on the state-of-the-art integrated corridor management strategies will be reviewed in Chapter 2, in which the existing literature and the summary of the ICM practice will be provided. Chapter 3 introduces a new maximum flow based signal control model to manage oversaturation on signalized arterials. The designed integrated control model will be discussed in Chapter 4 together with the simulation test results. Chapter 5 describes the implementation procedures of the SMART-Signal system on TH 55. In the end, concluding remarks are stated in Chapter 6.

Chapter 2. Review on the Theory and Practice of the Integrated Corridor Management

2.1 Background

Integrated control is not a new traffic management concept. The earliest research on integrated control at network level can be traced back to 1970s. Robertson and Vincent (1974) first consider the integration of bus schedules and signal timings. Later, van Aerde and Yagar (1988a) were the first to clearly address the importance of integrated control and discuss the required characteristics to operate an integrated control system. Van Aerde and Yagar (1988b) further propose a conceptual ICM approach, called INTEGRATION, which applies dynamic optimal route guidance as the control strategy in an integrated network. Their assumption of on-board driving information systems provides wide range applications by employing more sophisticated traffic controls.

After van Aerde and Yagar's work, researchers established various integrated traffic control frameworks, models, and solution approaches. The U.S. government noticed the potential benefits of ICM and started to document the research on this topic. In 2005, the U.S. Department of Transportation's (USDOT) Intelligent Transportation Systems (ITS) program launched the ICM Systems initiative (details are available at <http://www.its.dot.gov/icms/>), whose ultimate goal is "to provide the institutional guidance, operational capabilities, and ITS technology and technical methods needed for effective Integrated Corridor Management Systems". The initiative further propels the research on ICM in a regulated way, with the refined definitions for transportation corridor and ICM.

2.1.1 Definitions

With a broad definition of transportation corridor, tremendous studies on integrated control strategy can be classified into ICM. Papageorgiou (1995) systematically modeled ICM strategies, where a traffic corridor is simply defined as "a general highway network including both motorways and urban roads". Papageorgiou's definition is oriented to single-modal corridor control that is not sufficient to cover the broadly discussed ICM nowadays. A wider scope should be covered in the definition. In the ICM Program Plan, a succinct definition of transportation corridor is provided as "a combination of discrete parallel surface transportation networks (e.g., freeway, arterial, transit networks) that link the same major origins and destinations. It is defined operationally rather than geographically or organizationally". This definition emphasizes the operation aspect in a transportation corridor, and is more suitable in discussing ICM.

The simple definitions of corridor are not sophisticated enough to highlight the special characteristics and key elements of a corridor from the operational perspective of ICM. Responding to this need, Reiss et al. (2006) provide a refined definition of transportation corridor in their ICM initiative report:

Corridor—A largely linear geographic band defined by existing and forecasted travel patterns involving both people and goods. The corridor serves a particular travel market or markets that are affected by similar transportation needs and mobility issues. The corridor includes various networks (e.g., limited access facility, surface arterial(s), transit, bicycle, pedestrian pathway, waterway) that provide similar or complementary transportation functions. Additionally, the corridor includes cross-network connections that permit the individual networks to be readily accessible from each other.

After we clarify the study subject at corridor level, it is easier to categorize the ICM strategies. In early research, no clear definition of ICM was given. Van Aerde and Yagar (1988b) simply describe the integrated management as a way “to jointly optimize the combined network as an integrated unit”. In the ICM Program Plan, ICM is defined as “the coordination of individual network operations between adjacent facilities that creates an interconnected system capable of cross-network travel management.” To differentiate the ICM from other advanced traffic management systems, Reiss et al. (2006) report a refined definition of ICM as follows:

Integrated Corridor Management – ICM consists of the operational coordination of multiple transportation networks and cross-network connections comprising a corridor and the coordination of institutions responsible for corridor mobility. The goal of ICM is to improve mobility, safety, and other transportation objectives for travelers and goods. ICM may encompass several activities, for example:

- *Cooperative and integrated policy among stakeholders responsible for operations in the corridor.*
- *Concept of operations for corridor management.*
- *Improving the efficiency of cross-network junctions and interfaces.*
- *Mobility opportunities, including shifts to alternate routes and modes.*
- *Real-time traffic and transit monitoring.*
- *Real-time information distribution (including alternate networks).*
- *Congestion management (recurring and non-recurring).*
- *Incident management.*
- *Travel demand management.*
- *Public awareness programs.*
- *Transportation pricing and payment.*

The comprehensive definition of ICM summarizes the main attributes of ICM. It emphasizes the dependence and interaction between the activities imposed on sub-systems of a transportation corridor. The list of activities of ICM also provides a natural classification of ICM-related research topics in the literature.

2.1.2 Fundamentals

ICM Program Plan emphasizes three fundamental elements in implementing successful corridor management: institutional integration, operational integration, and technological integration. As noted in the ICM Systems initiative, the operational integration is the foundation of institutional and technological integrations. Thus, we will focus on reviewing the theoretical development in

operational integration, and briefly review the works on the institutional and technological integrations in practice.

From an operations perspective, a typical ICM system usually contains a set of information processing procedures and management processes, which support decision makers to identify appropriate controls. Since the role of information in ICM shifts to support the whole corridor rather than individual sub-systems, the information processing procedures in ICM need to integrate more components than those in individual management sub-systems, especially when more sub-systems are involved. For example, it may need to gather travelers' choice and their transition information between different transportation modes. Gathering reliable data is in essence for analyzing and delivering the underlying information to management processes (Cronin et al. 2010). Current data collection and information processing procedures in ICM heavily depend on applications of advanced information and communication technologies. Therefore, the advance of information processing modules in ICM is (highly) correlated with the advance of ITS technologies. With quick advances in communication technologies, an ICM system is able to archive robust and reliable data by low-cost data collection methodologies. However, the current advanced information has not been effectively used and integrated for ICM, mainly due to the absence of reliable forecasting methods and decision-support methodologies (based on existing data). There is a strong need for developing forecast and management architectures for ICM based on existing advanced information and communication technologies. Responding to this need, tremendous research efforts on ICM have been put on the development of various management methodologies in ICM. Due to the complexness and the broadness in composition of ICM system, the proposed methodologies are constructed from various perspectives.

2.2 Model Development

In current practice, existing ICM systems rely on reliable well-calibrated models (Cronin et al. 2010). Significant research effort has been paid to the analysis, modeling, and simulation methodologies. Based on the focus of those models, existing ICM models can be roughly categorized into two groups. The first group mainly focuses on the information provision and travelers' response, such as providing travel time information of different routes through Variable Message Sign (VMS), and the models are more macroscopic; while the second group emphasizes the traffic evolution and interaction, which is more microscopic.

2.2.1 Focus on information provision and travelers' response

Route guidance via certain media, such as Variable Message Sign (VMS) or on-board navigation system, is considered an effective way to advise motorists to better use network capacities. This group of studies was first conducted by Papageorgiou (1990), in which a macroscopic modeling framework for dynamic modeling and control of traffic networks was presented under time-varying demand conditions. In this model, the traffic network may include both freeways and urban roads and the control measures include individual and/or collective route guidance, signal setting and ramp metering. This approach was extensively studied and extended by the following researchers: Hawas & Mahmassani (1995) developed a procedure for real-time route guidance in congested vehicular traffic networks. Their approach collects information from a set of local

controllers scattered or distributed in the network and utilizes this information to guide vehicles. Messmer & Papageorgiou (1995) proposed a nonlinear optimization method to control motorway networks via Variable Message Signs (VMS). The problem is formulated as a dynamic, nonlinear, discrete-time optimal control problem with constrained control variables and can be solved by gradient based search methods. Ben-Akiva et. al (1997), used the DynaMIT (Dynamic Network Assignment for the Management of Information to Travelers) to generate real-time prediction-based guidance information. Pavlis & Papageorgiou (1999) and Minciardi (2001) further developed a simple decentralized feedback strategy for route guidance in traffic networks, where the measurable instantaneous travel times were used to generate control decisions. The objective of this type of models is to minimize the difference among the travel times from each origin node to all the possible destinations through the available routes in the network. Similar research work can be found at Yang and Yagar, 1995, Abdelghany et al., 1999, Mahmassani, 2001, Adler & Blue, 2002, Wang and Papageorgiou, 2002, Hamdar et al., 2006 and etc.

With the provided route guidance, different drivers may react differently. In essence, motorists' choices are the most important parts to determine eventual performance. To accommodate that there are a group of studies focusing on driver compliance. Peeta et al. (2000) investigated the effect of different message contents on driver response under VMS. The analysis was done through an on-site stated preference user survey and Logit models were developed for drivers' diversion decisions. The result shows that content in terms of the level of detail of relevant information significantly affects drivers' willingness to divert. Similarly, Kattan et al. (2010) conducted a survey of 500 Deerfoot Trail commuters in Calgary, Canada to examine the factors affecting drivers' compliance with VMSs. The results show that, 63.3% of drivers alter their trip plans with the information provided, comparing with 36.7% of drivers who did not alter their route. Regarding compliance model development, Peeta and Gedela (2001) proposed a VMS control heuristic framework, which ensures consistency with driver diversion response behavior. More recently, Paz and Peeta (2009) developed a fuzzy control modeling approach to determine the associated behavior-consistent information-based network control strategies. The approach can provide more robust performance compared to the standard user or system optimal information strategies. Lee et al. (2010) explored the factors affecting alternative route choices of car drivers with VMSs by adopting a method called LOTUS. The study pointed out that travel-time saving is not the single dominant factor for driver route choice under information provision.

2.2.2 Focus on traffic evolution and interaction

The traffic evolution and interaction between sub-systems also attracted many researchers' attention. In 1993, Chang et al. (1993) presented a dynamic system-optimal control model (DSOCM) for commuting corridors which consist of both freeway and surface streets. The proposed model considers the complex interactions among the freeway, surface street and diversion flows. Optimal time-dependent ramp metering rates and signal settings can be obtained by solving the model. Stephanedes and Kwon (1993) introduced an adaptive demand-diversion predictor which specifically considers the influence of traffic diversion to ramp metering and intersection signal timings. Following that, Papageorgiou (1995) developed an integrated control approach for traffic corridors including both freeways and signalized arterials based on the store-and-forward modeling philosophy. The control objective is to minimize the total delay or the

total time spent in the network. The formulated optimal-control problem may provide traffic-responsive queue management, particularly under saturated traffic conditions. Later, Wu and Chang (1999) proposed a control model in an on-line environment which integrates ramp metering, intersection signal timing and off-ramp diversion under non-recurrent congestion. The approach models traffic state evolution on surface streets and estimate time-dependent model parameters adaptively with real-time traffic measurements. Kotsialos et al. (2002) proposed a generic formulation in the format of discrete-time optimal control problem and it can be solved by a feasible-direction algorithm. More recently, Liu and Chang (2010) and Liu et al. (2011) introduced a multi-objective optimization model to maximize the utilization of the available corridor capacity. Because of model complexity, the optimal diversion rates can be obtained through a genetic algorithm-based technique. This group of models specifically considers the queue formation and dispersion among different systems, however, the complexity of formulation largely limits the practical implementation.

2.3 Operations

In this section, we will review those ICM strategies which are implemented or being tested in operation, as well as the lessons learned from field implementations. ITS aims to use advanced technologies to improve transportation on many levels, such as reduce congestion, enhance safety, mitigate the environmental impacts of transportation systems, enhance energy performance, and improve productivity (Sussman et al., 2000). A brief review on ITS technologies of existing transportation infrastructures will be first given in the following.

2.3.1 ITS Technologies

As stated, the operation of Integrated Corridor Management (ICM) highly depends on the development of Intelligent Transportation System (ITS) because there is a lot of information processing and information sharing between different sub-systems. The integration of different ITS infrastructures just constructs the base for various ICM control strategies.

The ITS in arterial management mainly includes two areas, the adaptive control strategies (ACS) and the advanced traveler information system (ATIS). The ACS optimizes intersection signal timing plans in real time, based on current traffic conditions and demand. Representative ACS systems include SCOOT (Hunt et al., 1982), OPAC (Gartner, 1983), RHODES (Sen & Head, 1997) and etc. These adaptive control system works great under light and medium traffic conditions, however, the performance deteriorates in case of saturated traffic conditions. ATIS for arterials provides information on arterial conditions (e.g., travel speeds, travel time, incidents) to travelers through certain media.

On freeway management, ramp metering is a major and highly effective management tool, which can reduce the traffic congestion on freeway (Sussman et al., 2000). These metering strategies can be divided into two groups, i.e. fixed time strategies, such as Wattleworth, 1965 and reactive metering strategies, such as Papageorgiou et al., 1998. These strategies have been extensively evaluated not only in simulation but also in field (Chang and Stephanedes, 1993, Haj-Salem and Papageorgiou, 1995, Zhang and Recker, 1999, Hasan and Ben-Akiva, 2003 Tian, 2007 and Ahn et al., 2007). Variable speed limits and dynamic lane controls continue to show promise, but are

not yet widely deployed in the United States. Variable Message Sign (VMS) is a widely used tool to provide information to drivers, allowing drivers to re-route before arriving to the network bottleneck.

2.3.2 Practices and Pioneer Sites

In practice, the application of the ICM concept is still at a very early stage. Because different transportation infrastructures usually belong to different agencies, communication and policy barriers have restricted most of the ICM application within the same traffic mode. Most of the work related to cross-modal integration still remains at the policy research level (Alm et al., 2008), such as cost-benefit analysis, incentive analysis and agreement analysis, and the evaluations are still based on simulation studies (Alexiadis, 2008).

From fall 1994 through spring 1999 in Irvine, California, a systematic evaluation of the performance and effectiveness of an integrated corridor-level adaptive control system was attempted. The California Department of Transportation (Caltrans), the city of Irvine, and two private-sector consultants were involved in the field test. Because of the failure of any of the planned technologies to be successfully implemented in the field, the test failed to provide a technical evaluation on the integrated system. However, two most valuable lessons were learned through the process: “It is important to incorporate detailed technical specifications in contract documents” and “there is a strong need for complete technical review and an appropriate level of technical understanding on the part of the contracting agency.” (MacCarley et al., 2002) Later, a new integrated traffic-responsive urban corridor control strategy IN-TUC (integrated traffic-responsive urban control) was developed and applied to the M8 corridor network in Glasgow, Scotland, which includes signal control, ramp metering, and VMS control. The results of the preliminary simulation investigations as well as the results of the field implementation and evaluation of the strategy seem promising (Diakaki et al., 2000). More recently, Perugu et al. 2007 developed an Integrated Corridor Control system with Access management (IUCC-ACCESS). The heuristic algorithm integrates SCOOT algorithm, and BOTTLENECK-access metering algorithm with dynamic rerouting for selected time horizon. , a simulation-based test bed was built to test the proposed model and the test results showed the effectiveness of improve network performance.

In 2005, the U.S. Department of Transportation’s (USDOT) Intelligent Transportation Systems (ITS) program launched the ICM Systems initiative and eight pioneer sites were selected. The eight sites are located in Oakland and San Diego, CA; Dallas, Houston, and San Antonio, TX; Montgomery County, MD; Seattle, WA; and Minneapolis, MN. These pioneer sites are all recognized leaders in the area of congestion management and the corresponding corridors are already equipped with advance ITS infrastructures, such as HOV/HOT lanes, real time arterial signal and ramp metering control, rapid transit services and etc. (FHWA, 2007). All eight sites participated in the ICM Initiative's initial phase (concepts of operations and system requirement), which was completed in 2007 (FHWA, 2007). These sites have developed multimodal ICM strategies that apply new institutional and operational approaches and advanced technologies to existing infrastructure. In late 2009, Dallas, TX and San Diego, CA were selected by the U.S. DOT to demonstrate their ICM systems. Dallas will integrate the regional systems and operations along the US-75 corridor using a decentralized approach. Travelers will have access

to real-time information about traffic and travel times, public transit, and parking availability through wireless and web-based alerts as well as dynamic message signs on the roads (FHWA, 2008a). San Diego aims to proactively and collaboratively manage the I-15 corridor to maximize transportation system performance. Through collaboration among the corridor's institutional partners, the system will enable travelers the opportunity to make convenient shifts among modes and routes (FHWA, 2008b).

2.4 Challenges and Opportunities

2.4.1 On the Modeling Side

The development of ICM models has two major challenges: the origin – destination information estimation and drivers' compliance rate estimation. Most existing models require the O-D matrix as inputs to generate specific control decisions, however, as well known, the estimation of network O-D information is extremely difficult, especially under congested scenarios. How to successfully get time-dependent network O-D information will still be a challenging task in the ICM domain, since it determines the possibility of practical implementation for most developed models. On the other hand, even if the optimal control strategies can be generated, motorists' reaction to these control decisions is essentially the key to the final network performance. A more robust and accurate real-time estimation model is desired to estimate drivers' compliance rates to given control strategies.

2.4.2 On the Operation Side

The implementation of ICM control strategies is mainly restricted by two factors: the lack of real time traffic data collection and the communication barriers between different agencies. The problem of real time data collection becomes more prominent in the area of arterial management, due to the absence of ITS infrastructures. More implementation of advanced data collection system, which can provides the most updated traffic information (travel time, speed, incident and etc.), is still an important step to achieve ICM goals. The communication barriers between transportation agencies (arterial, freeway, transit and etc.), due to either policy limitation or technical problem, still largely undermine the base of ICM. The construction of a central system, which can integrate all the real time traffic information between different agencies and dispatch the optimal control strategies to sub-systems, still needs a huge effort from various departments.

Further, through field implementation evaluations, we can see that the considerations on ICM should go far beyond quantitative cost-benefit analysis in proposed strategies and infrastructure investment. Some qualitative factors, such as enhancing intermodal transition reliability, improving freight mobility, or reducing environmental impacts to better support livability and sustainability, should be considered in the future as well.

2.5 Summary

In general, the Integrated Corridor Management (ICM) is a comprehensive and challenging problem. It consists of the operational coordination of multiple transportation networks (freeway,

arterial, transit, bicycle, pedestrian pathway and etc.) and cross-network connections. Due to the complexity and the broadness of the ICM system, control models are constructed from various perspectives. Based on the focus of those models, existing ICM models can be roughly categorized into two groups. The first group mainly focuses on the information provision and travelers' response (more macroscopic) and the second group emphasizes the traffic evolution and interaction (more microscopic). How to accurately estimate the origin – destination information and drivers' compliance rate in real time is the key problem waiting to be solved for existing ICM models. In practice, due to the communication and policy barriers between different agencies, the application of the ICM concept is still at the very early stage. Most work remains at the policy research level and the evaluations are still based on simulation studies. In 2005, the USDOT Intelligent Transportation Systems (ITS) program launched the ICM Systems initiative and eight pioneer sites were selected to generate the concept of ICM operation. Later, Dallas, TX and San Diego, CA were selected to demonstrate their ICM systems. Limited field experience has already shown that, in order to successfully implement ICM, different transportation agencies must cooperate with each other and share resources and real time traffic information. Eventually, the integrated system will improve transportation efficiency, robustness and flexibility, which is very promising, especially under the increasing trend of traffic demand.

Chapter 3. Managing Oversaturated Signalized Arterials: A Maximum Flow Based Approach

3.1 OSI-based Mitigation Strategies

The proposed model is built upon the recent development in the diagnosis of oversaturation (Wu et al., 2010) using high-resolution vehicle actuation and signal event data collected by the SMART-Signal system (Systematic Monitoring of Arterial Road Traffic and Signals) (Liu & Ma, 2009; Liu et al., 2010). In Wu et al. 2010, an Oversaturation Severity Index was proposed to quantify the severity level of oversaturation by measuring its detrimental effects. Detrimental effect is characterized by either a residual queue at the end of a cycle or a spillover from downstream traffic, both of which create “unusable” green time. In the case of residual queue, the “unusable” green time is the equivalent green time to discharge the residual queue in the following cycle, but for spillover, the “unusable” green time is the time period during which an downstream link is blocked and therefore the discharge rate is zero. OSI is defined as the ratio between unusable green time and total available green time in a cycle, which is a non-negative percentage value between 0 and 100, with 0 indicating no detrimental effect for signal operation, and 100 being the worst that all available green time becomes unusable. OSI is further differentiated into *TOSI* (Temporal Oversaturation Severity Index, caused by the residual queue that creates the detrimental effect in temporal dimension) and *SOSI* (Spatial Oversaturation Severity Index, caused by the spillover that creates the detrimental effect in spatial dimension). Specifically, *TOSI* and *SOSI* can be calculated by Eq. (3.1). More detailed derivation and explanation of the two indices can be found in Wu et al. (2010).

$$\left\{ \begin{array}{l} TOSI = \frac{\text{Green time to discharge residual queue}}{\text{Total available green time}} \\ SOSI = \frac{\text{Unusable green time due to spillover}}{\text{Total available green time}} \end{array} \right. \quad \text{Eq. (3.1)}$$

With *TOSI* and *SOSI*, not only can the severity level of oversaturation be quantified, but also the causes of arterial traffic congestion can be identified. Positive *TOSI* indicates that the available green time is insufficient for queue discharge and a residual queue is formed at the end of a cycle; and positive *SOSI* indicates that the queue length at the downstream link has reached the upstream intersection and blocked the discharging traffic. Based on measured *TOSI* and *SOSI* values, three basic mitigation strategies are designed for different oversaturation scenarios between two intersections.

1) Green Extension for Scenario 1: $TOSI > 0$ & $SOSI = 0$.

Since a positive *TOSI* value indicates a residual queue at the end of a cycle and zero *SOSI* value indicates that there is still spare capacity to store vehicles in the downstream link, the strategy to deal with this situation is to extend the green time for the oversaturated phase.

Figure 3.1(a) illustrates this case by presenting the shockwave profiles for two intersections. After extending the green in Figure 3.1 (b), the residual queue disappears and $TOSI$ becomes zero. The green extension can be calculated as the following (Eq. (3.2)).

$$\Delta g_{n,i} = TOSI_{n,i} \times g_{n,i} \quad \text{Eq. (3.2)}$$

where $\Delta g_{n,i}$ is the adjustment to the green time at intersection n for phase i ; $TOSI_{n,i}$ is the $TOSI$ value at intersection n for phase i ; and $g_{n,i}$ is the green time at intersection n for phase i . Note that positive $\Delta g_{n,i}$ means green extension; and a negative value means green reduction. By extending green, the start time of the following red signal will be postponed.

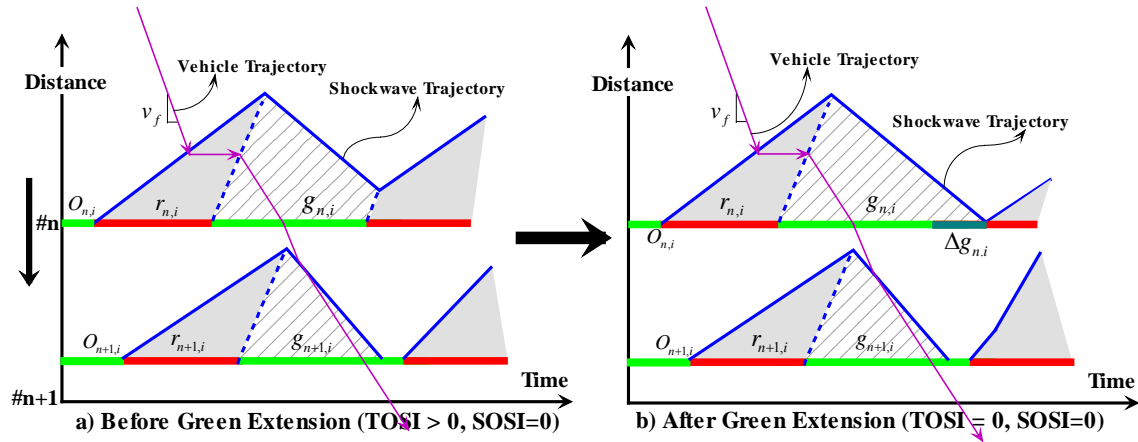


Figure 3.1 Green extension for Scenario 1

2) Red Extension for Scenario 2: $TOSI = 0$ & $SOSI > 0$.

If $SOSI$ is larger than zero, it indicates that the downstream queue spills back to the upstream intersection and results in unusable green time as shown in Figure 3.2 (a). But since $TOSI$ is zero, all queued vehicles can be discharged even with reduced green time. One way to remove downstream spillover is to gate the upstream flow by extending the red time. The red extension can be calculated as the following (Eq. (3.3)).

$$\Delta r_{n,i} = SOSI_{n,i} \times g_{n,i} \quad \text{Eq. (3.3)}$$

Where $\Delta r_{n,i}$ is the adjustment to the red time at intersection n for phase i ; and $SOSI_{n,i}$ is the $SOSI$ value at intersection n for phase i . The positive $\Delta r_{n,i}$ means red extension and a negative value means red reduction. Note that by extending the red time, the start of the following green will be postponed.

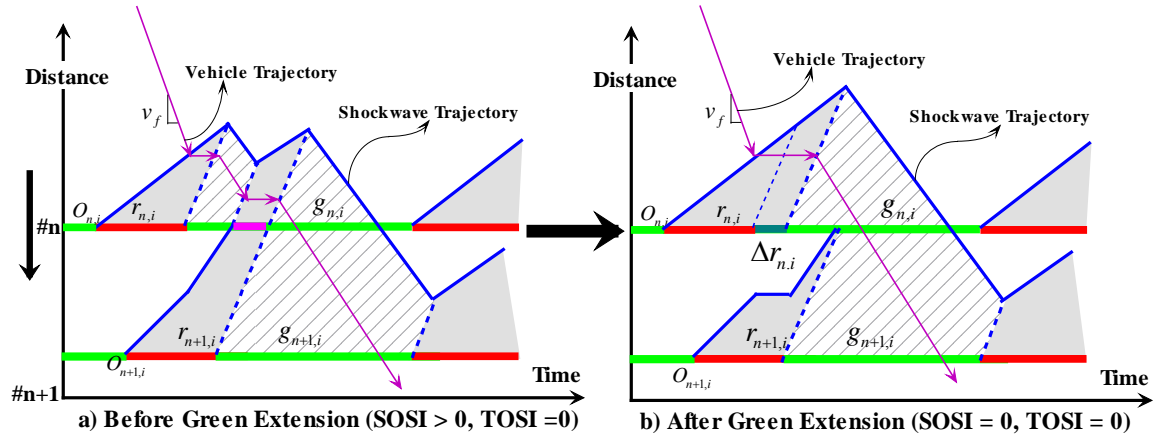


Figure 3.2 Red extension for Scenario 2

3) Downstream Red Reduction for Scenario 3: $TOSI > 0$ & $SOSI > 0$.

A more serious situation exists when both $TOSI$ and $SOSI$ are larger than zero, as shown in Fig. 3(a). In this case, at the upstream intersection a portion of the green time is unused because of the downstream spillover. At the same time, the useable green time at the upstream intersection is not sufficient to discharge queued vehicles, i.e., a residual queue exists. One way to deal with this scenario is to increase downstream capacity by reducing the red time at the downstream intersection. As shown in Figure 3.3, by reducing the downstream red, positive $TOSI$ and $SOSI$ values for the upstream intersection will be reduced. Once the downstream spillover is removed or reduced, the unusable green time at the upstream intersection may become available and can be used to discharge the residual queue. If $TOSI < SOSI$, the residual queue can be cleared by using this strategy. The reduction of downstream red can be calculated as the following (Eq. (3.4)).

$$\Delta r_{n+1,i} = -SOSI_{n,i} \times g_{n,i} \quad \text{Eq. (3.4)}$$

It should be noted, as an alternative to reduce the downstream red, we can also deal with situation by combining the methods for scenario 1 and 2 together, i.e., extending both the red and green times.

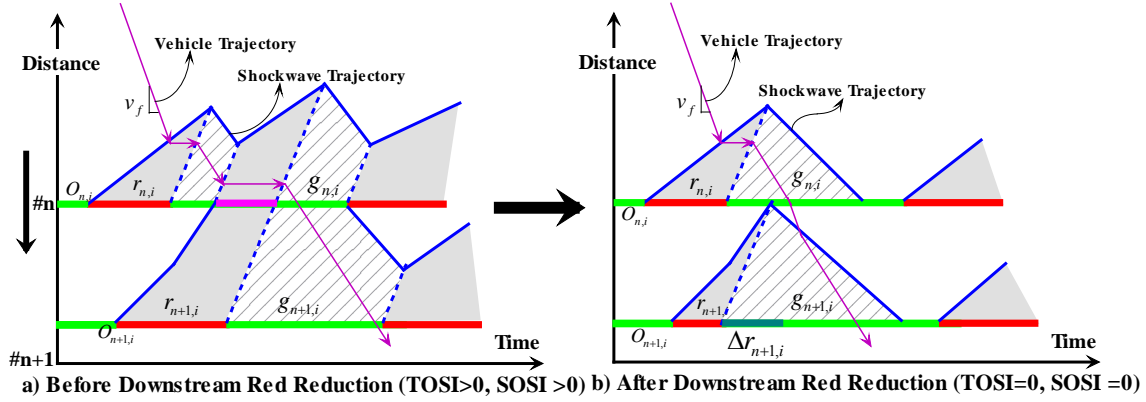


Figure 3.3 Red reduction (at downstream intersection) for Scenario 3

Among the three strategies, extending green (strategy 1) is to increase the discharge capacity for the oversaturated phase; extending red (strategy 2) is to gate traffic arrivals at the upstream intersection; and reducing downstream red (strategy 3) is to remove the downstream bottleneck by discharging the queue earlier at downstream intersection. By considering maximum/minimum green and storage space limitations on side streets, the strategies introduced above may be directly applied for an isolated intersection or two intersections in tandem. However, if multiple intersections along a route have oversaturated problems, a systematic strategy is needed for two reasons. First, the increase of green time of an upstream approach may create oversaturation on the downstream link and secondly, capacity constraints at a downstream phase may limit the possible signal timing adjustments for the upstream phase.

3.2 Problem Definition and Assumptions

The proposed model aims to solve the signal control problem when multiple intersections along a route are oversaturated. This route is defined as an oversaturated route, see Figure 3.4. An oversaturated route may be a straight line (Figure 3.4 a) or may not (Figure 3.4 b), since the oversaturation for some intersections may be caused by the turning movements. Given the $TOSI$ and $SOSI$ values for each movement, it is easy to identify oversaturated routes, since all intersections along this route will have positive $TOSI$ and/or $SOSI$ values. At the start of control period t , the model determines the signal timing changes according to the average $TOSI$ and $SOSI$ values at the immediate past control period $t-1$, i.e., $TOSI_{n,i}(t-1)$, $SOSI_{n,i}(t-1)$. One control period usually has 3~4 cycles.

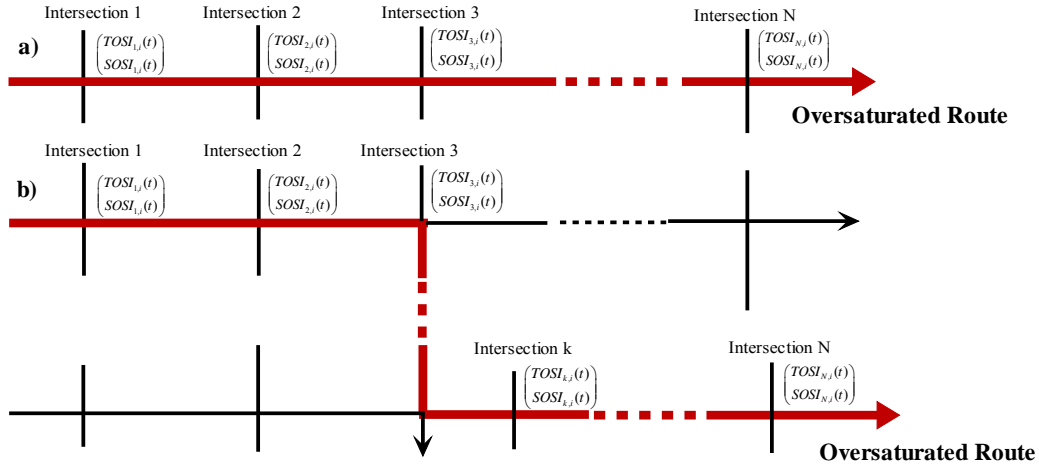


Figure 3.4 Intersections along an oversaturated route

Two important assumptions need to be clarified. First, we assume that cycle lengths among all intersections along an oversaturated route keep unchanged. Secondly, the traffic condition in control period $t+1$ will be the same as that in control period t , if the signal timing parameters (e.g., offset, green duration, and cycle) keep the same. It is a reasonable assumption because traffic conditions usually do not change significantly within a short time period.

3.3 The Maximum-Flow Program (MFP)

Based on the three basic mitigation strategies, an optimization model is proposed. It aims to mitigate the oversaturated conditions at signalized arterial/network by maximizing the discharging capacity.

3.3.1 Control Variables

In the maximum flow model, two sets of control variables $\Delta r_{n,i}(t)$ and $\Delta g_{n,i}(t)$, namely red time changes and green time changes for phase i at intersection n , are introduced for each oversaturated phase. The two control variables have direct association with specific oversaturation mitigation strategies. Whether to change red or green is determined by the causes of the oversaturation. Changing red times (i.e. $\Delta r_{n,i}$) aims to eliminate spillover; and changing green times (i.e. $\Delta g_{n,i}$) aims to clear residual queues. A positive red time change (**red extension**) means that extra red time is added. Since the cycle length is kept unchanged, the green start would be postponed with the red extension (see Figure 3.5 a) and the total green time is reduced. A negative red time change (**red reduction**) means a portion of red time is cut from the end of red, therefore, green start will be advanced (see Figure 3.5 b) and the total green time is increased. Similarly, a positive green time change (**green extension**) indicates that additional green time is added to the original end of the green time (see Figure 3.5 c), and a negative green time change (**green reduction**) represents that some green time is cut from the end of green (see

Figure 3.5 d). Depending on the offset reference point used for the intersection (start of yellow, start of green, barrier crossing, etc.), each case of adjusting green or red may require a corresponding change to the offset and green split values.

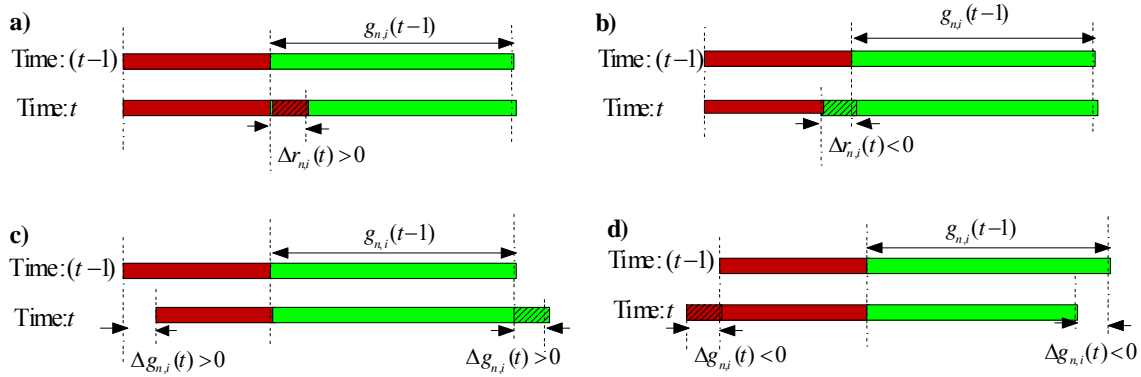


Figure 3.5 Red time changes & green time changes

The values $\Delta r_{n,i}(t)$ and $\Delta g_{n,i}(t)$ can be easily transformed into the values of new offset and green duration, which can easily be modified in the signal timing plan. If we assume that the oversaturated phase is the coordinated direction and the green start time of the coordinated phase is the offset reference point, Eq. (3.5) can be used to calculate the new offset and green duration after adjustment, where $o_n(t)$ is the offset value at time t and c_n is the cycle length for intersection n . Figure 3.6 presents an example of signal timing changes at one intersection with $\Delta r_{n,i}(t) < 0$ and $\Delta g_{n,i}(t) > 0$. Note that, if there is no change on the offset, the designed signal timing can be achieved in the immediately next cycle; however, if there is a change on the offset, a transition period (1~2 cycles) is unavoidable in order to get to the new timing.

$$\begin{cases} o_n(t) = o_n(t-1) + \Delta r_{n,i}(t) \\ g_{n,i}(t) = g_{n,i}(t-1) - \Delta r_{n,i}(t) + \Delta g_{n,i}(t) \\ r_{n,i}(t) = c_n - g_{n,i}(t) \end{cases} \quad \text{Eq. (3.5)}$$

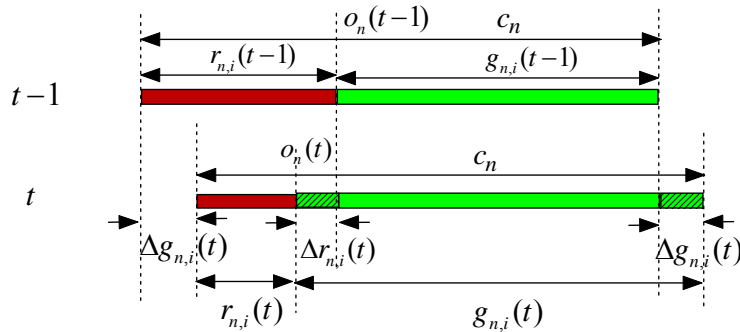


Figure 3.6 Signal timing changes ($\Delta r_{n,i}(t) < 0$, $\Delta g_{n,i}(t) > 0$)

3.3.2 Constraint Analysis

In the following, for the sake of simplicity, we use $S_{n,i}(t)$ to represent the unusable green time caused by spillover (i.e., $SOSI > 0$) at intersection n for phase i during time period of t , and use $T_{n,i}(t)$ to represent the unusable green time caused by residual queue (i.e., $TOSI > 0$).

$$\begin{cases} S_{n,i}(t) = SOSI_{n,i}(t) \times g_{n,i}(t) \\ T_{n,i}(t) = TOSI_{n,i}(t) \times g_{n,i}(t) \end{cases} \quad \text{Eq. (3.6)}$$

(1) Spillover Elimination

The proposed control model aims to eliminate spillover between intersections. In order to eliminate the spillover at intersection n , one can either extend the red time at the current intersection n (i.e., apply gating at the upstream intersection, see Figure 3.2, or reduce the red time at the downstream intersection $n+1$ (i.e., discharge the downstream queue earlier, see Figure 3.3), or a combination of the two strategies. As described in Figure 3.7, extending the red at intersection n by $\Delta r_{n,i}(t)$ ($\Delta r_{n,i}(t) > 0$) will make the unusable green time caused by spillover shorter by $\Delta r_{n,i}(t)$; On the other hand, reducing the red at intersection $n+1$ by $|\Delta r_{n+1,i}(t)|$ ($\Delta r_{n+1,i}(t) < 0$) will make the unusable green caused by spillover at intersection n shorter by $|\Delta r_{n+1,i}(t)|$. Therefore, in order to eliminate spillover at intersection n , the difference of red time changes between intersection n and intersection $n+1$ should be equal to the unusable green time caused by spillover at intersection n , i.e., $S_{n,i}(t-1)$, see Eq. (3.7).

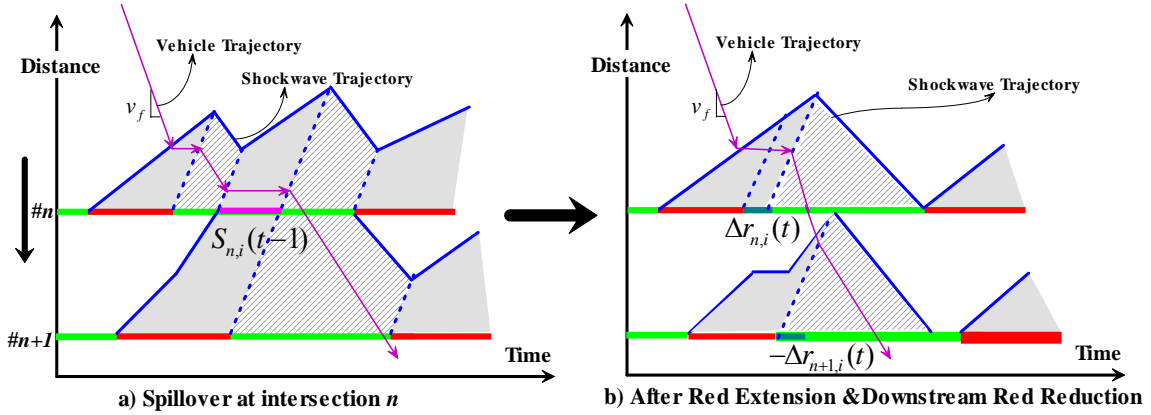


Figure 3.7 An example of applying strategies 2 & 3 to eliminate spillover

$$\Delta r_{n,i}(t) - \Delta r_{n+1,i}(t) = S_{n,i}(t-1), \forall n \in \{1, \dots, N-1\} \quad \text{Eq. (3.7)}$$

(2) Residual Queue Elimination

If Eq. (3.7) is satisfied, the spillovers are supposed to be eliminated during control period t . Then the green time change $\Delta g_{n,i}(t)$ for each intersection is used to eliminate residual queue. If the red time and green time changes at intersection n are $\Delta r_{n,i}(t)$ and $\Delta g_{n,i}(t)$ respectively, the total green time at intersection n for control period t would be $[-\Delta r_{n,i}(t) + \Delta g_{n,i}(t) + g_{n,i}(t-1)]$. If

Intersection $n+1$ has residual queue in control period $t-1$ and the corresponding unusable green time is $T_{n+1,i}(t-1)$, in order to eliminate residual queue of Intersection $n+1$ at control period t , the difference of total green time between Intersection $n+1$ and its upstream intersection n should be equal to $T_{n+1,i}(t-1)$, i.e., Eq. (3.8) should hold.

$$\begin{aligned} & [\Delta g_{n+1,i}(t) - \Delta r_{n+1,i}(t) + g_{n+1,i}(t-1)] - [\Delta g_{n,i}(t) - \Delta r_{n,i}(t) + g_{n,i}(t-1)] \\ & = T_{n+1,i}(t-1), n \in \{1, \dots, N-1\} \end{aligned} \quad \text{Eq. (3.8)}$$

Substitute Eq. (3.7) into Eq. (3.8),

$$\begin{aligned} & \Delta g_{n+1,i}(t) - \Delta g_{n,i}(t) \\ & = T_{n+1,i}(t-1) - S_{n,i}(t-1) - [g_{n+1,i}(t-1) - g_{n,i}(t-1)], n \in \{1, \dots, N-1\} \end{aligned} \quad \text{Eq. (3.9)}$$

(3) Available Green Constraints

For each intersection along the oversaturated route, the green time increase at control period t , i.e., $\Delta g_{n,i}(t) - \Delta r_{n,i}(t)$ is constrained by the available green time $g_{n,i}^a(t)$ for intersection n and phase i , see Eq. (3.10).

$$\Delta g_{n,i}(t) - \Delta r_{n,i}(t) \leq g_{n,i}^a(t), \forall n \in \{1, \dots, N\} \quad \text{Eq. (3.10)}$$

If $Z_{n,i}$ is the set of conflicting phases to phase i at intersection n , the available green time $g_{n,i}^a(t)$ can be computed by considering the maximum queue size for each of these conflicting phases in the immediate past control interval $t-1$, see Eq. (3.11). Here c_n is the cycle length for intersection n , $q_{n,p}^{\max}(t-1)$ is the maximum queue size per lane for phase p at intersection n at control interval $t-1$ and $s_{n,p}$ is the saturation flow rate per lane for phase p of intersection n . $q_{n,p}^{\max}(t-1) / s_{n,p}$ calculates how much green time is needed to discharge the queue of $q_{n,p}^{\max}(t-1)$. α is a weighting term, which represents users' perspective on the importance of queues on conflicting phases when calculating the available green for oversaturated phase i . When the maximum queue length for phase p , i.e. $q_{n,p}^{\max}(t-1)l$ (where l is the jammed space headway), is shorter than the corresponding link length $L_{n,p}$, we would like to only account for a portion

($\alpha = \beta, \beta \in (0,1)$) of these queues because we want to maximize the discharging capacity for the oversaturated route to reduce congestion; however, if the maximum queue length for phase p is already longer than the link length, all the queues need to be considered ($\alpha = 1$), otherwise these queues will block further upstream intersections. One should note that, the smaller the β is, the more expected extra capacity will be assigned to the oversaturated route, the faster the queues on conflicting phases will grow and the more delay will be introduced to conflicting phases. A recommended value for β would be around 0.5.

$$g_{n,i}^a(t) = c_n - \sum_{p \in Z_{n,i}} \alpha \left[q_{n,p}^{\max}(t-1) / s_{n,p} \right] - g_{n,i}(t-1) \quad \text{Eq. (3.11)}$$

$$\text{Where } \alpha = \begin{cases} \beta, & \text{if } q_{n,p}^{\max}(t-1)l < L_{n,p} \\ 1, & \text{if } q_{n,p}^{\max}(t-1)l \geq L_{n,p} \end{cases}$$

3.3.3 Maximum Flow Based Control Model

(1) One oversaturated route

The objective of the control model is to maximize the discharging capacity along the oversaturated route. At each control period t , it is equivalent to maximizing the total green time at the first intersection of the route, i.e., $(\Delta g_{1,i}(t) - \Delta r_{1,i}(t) + g_{1,i}(t-1))$. Since $g_{1,i}(t-1)$ is the green time during control period $t-1$, at the start of control period t , maximizing $(\Delta g_{1,i}(t) - \Delta r_{1,i}(t) + g_{1,i}(t-1))$ is equivalent to maximizing $(\Delta g_{1,i}(t) - \Delta r_{1,i}(t))$. Therefore, the complete control model can be expressed in Eq. (3.12). The first and second constraints ensure the elimination of spillover and residual queues between intersections and the third constraint considers the available green time. Note that, if the green duration for the oversaturated route is shortened at some intersection due to red extension or green reduction, the non-oversaturated directions will receive more green time and that may cause more vehicles from the non-oversaturated directions to be added to the oversaturated route because of turning. However, in most cases, the traffic volume discharged from the non-oversaturated directions is much smaller than the one from the oversaturated direction, so the strategy will not worsen the traffic condition. Further, if this strategy makes the previously non-oversaturated route become oversaturated, the traffic condition will then be considered as the case with two intersecting oversaturated routes, as will be discussed in the next section.

$$\max \Delta g_{1,i}(t) - \Delta r_{1,i}(t)$$

s.t.

$$\Delta r_{n,i}(t) - \Delta r_{n+1,i}(t) = S_{n,i}(t-1), \quad n \in \{1, \dots, N-1\}$$

$$\Delta g_{n+1,i}(t) - \Delta g_{n,i}(t) \quad \text{Eq. (3.12)}$$

$$= T_{n+1,i}(t-1) - S_{n,i}(t-1) - [g_{n+1,i}(t-1) - g_{n,i}(t-1)], \quad n \in \{1, \dots, N-1\}$$

$$\Delta g_{n,i}(t) - \Delta r_{n,i}(t) \leq c_n - \sum_{p \in Z_{n,i}} \alpha [q_{n,p}^{\max}(t-1) / s_{n,p}] - g_{n,i}(t-1), \quad n \in \{1, \dots, N\}$$

If we treat the signal timing changes of each oversaturated phase along the route as flows in a network, the above linear program can be seen as a multi-commodity maximum flow problem. The corresponding network $G(M, A)$ is shown in Figure 3.8 with node set M and arc set A . Each intersection $n \in \{1, \dots, N\}$ along the oversaturated route is corresponding to a node n in the network. The node set M also includes two dummy nodes: the source node S_i and the sink node D_i . The arc coming out of node $n \in \{1, \dots, N\}$ has a capacity constraint $g_{n,i}^a(t)$, which limits the total flow it can carry. There are two kinds of flows or commodities in the network, namely red reduction $-\Delta r_{n,i}(t)$ and green extension $\Delta g_{n,i}(t)$. In Figure 3.8, we place the flow of red reduction on the upper side of the corresponding arc and the green extension on the lower side. The total flow of the two commodities on each arc is constrained by the arc capacity, i.e., $\Delta g_{n,i}(t) - \Delta r_{n,i}(t) \leq g_{n,i}^a(t)$, $\forall n \in \{1, \dots, N\}$. At each node $n \in \{2, 3, \dots, N\}$, there is an input flow $d_{n,i}^r(t)$ for red reduction and another input flow $d_{n,i}^g(t)$ for green extension. The definition of $d_{n,i}^r(t)$ and $d_{n,i}^g(t)$ is shown in Eq. (3.13). From a network flow point of view, at each node, the two input flows are external demands for the two commodities, respectively.

$$\begin{cases} d_{n+1,i}^r(t) = S_{n,i}(t-1) \\ d_{n+1,i}^g(t) = T_{n+1,i}(t-1) - S_{n,i}(t-1) - [g_{n+1,i}(t-1) - g_{n,i}(t-1)] \end{cases}, \quad n \in \{1, \dots, N-1\}$$

Eq. (3.13)

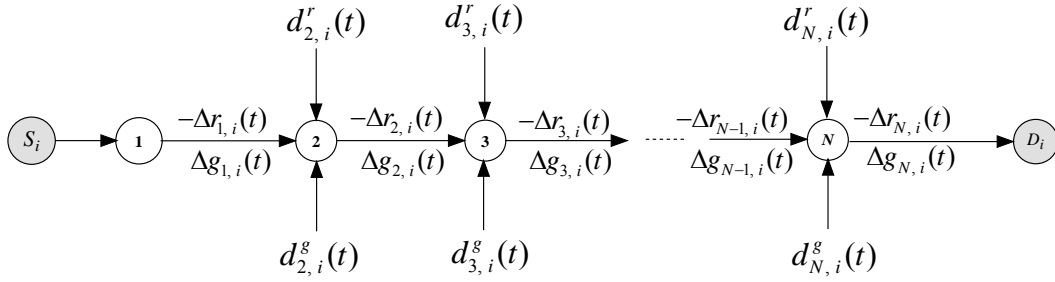


Figure 3.8 Maximum flow network for one oversaturated route

(2) Two intersecting oversaturated routes

If there are two intersecting oversaturated routes, similar approach can be applied to construct a maximum flow problem. As shown in Figure 3.9, Route 1 includes intersection 1, intersection 2, ..., intersection N , and follows the direction of phase i ; Route 2 includes intersection $1'$, intersection $2'$, ..., intersection N' , and follows the direction of phase j . Two oversaturated routes intersect with each other at the critical intersection I .

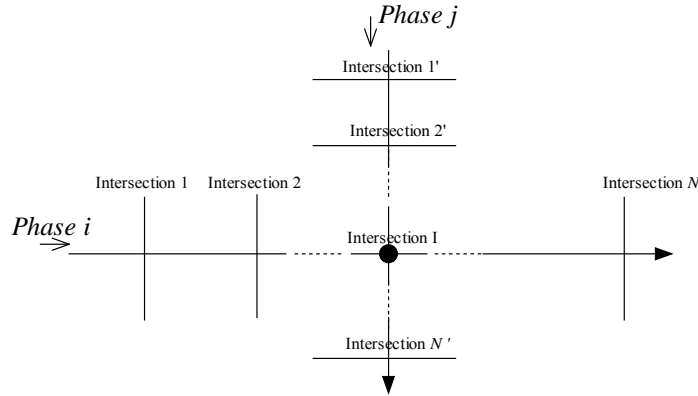


Figure 3.9 Two intersecting oversaturated routes

For two intersecting oversaturated routes, the objective of the control model is to maximize the total flows for both routes while satisfying spillover elimination, residual queue elimination and flow-capacity constraints. At intersection I , the available green for both phase i and j ($g_{I,i\&j}^a(t)$) can be calculated by Eq. (3.14), where $Z_{I,i\&j}$ is the set of conflict phases to phase i and j at intersection I .

$$g_{I,i\&j}^a(t) = c_n - \sum_{p \in Z_{I,i\&j}} \alpha \left[q_{n,p}^{\max}(t-1) / s_{n,p} \right] - g_{I,i}(t-1) - g_{I,j}(t-1) \quad \text{Eq. (3.14)}$$

The control model can be formulated in the form of Eq. (3.15). Most constraints have similar meanings as what we described for Eq. (3.12), except for the last one, which is the available green constraint at intersection I .

$$\begin{aligned}
& \max (\Delta g_{1,i}(t) - \Delta r_{1,i}(t)) + (\Delta g_{1',j}(t) - \Delta r_{1',j}(t)) \\
& s.t. \\
& \Delta r_{n,k}(t) - \Delta r_{n+1,k}(t) = S_{n,k}(t-1), \\
& \quad n \in \{1, \dots, N-1\} \& k = i \text{ or } n \in \{1', \dots, N'-1\} \& k = j \\
& \Delta g_{n+1,k}(t) - \Delta g_{n,k}(t) = T_{n+1,k}(t-1) - S_{n,k}(t-1) - [g_{n+1,k}(t-1) - g_{n,k}(t-1)], \\
& \quad n \in \{1, \dots, N-1\} \& k = i \text{ or } n \in \{1', \dots, N'-1\} \& k = j \\
& \Delta g_{n,k}(t) - \Delta r_{n,k}(t) \leq c_n - \sum_{p \in Z_{n,k}} \alpha [q_{n,p}^{\max}(t-1) / s_{n,p}] - g_{n,k}(t-1), \\
& \quad n \in \{1, \dots, I-1\} \cup \{I+1, \dots, N\} \& k = i \\
& \quad \text{or } n \in \{1', \dots, (I-1)'\} \cup \{(I+1)', \dots, N'\} \& k = j \\
& -\Delta r_{I,i}(t) - \Delta r_{I,j}(t) + \Delta g_{I,i}(t) + \Delta g_{I,j}(t) \\
& \quad \leq c_n - \sum_{p \in Z_{I,i \& j}} \alpha [q_{n,p}^{\max}(t-1) / s_{n,p}] - g_{I,i}(t-1) - g_{I,j}(t-1)
\end{aligned} \tag{3.15}$$

The corresponding maximum flow network is presented in Figure 3.10. There are two source nodes S_i and S_j , and two sink nodes D_i and D_j , for two oversaturated routes respectively. For each direction, it has similar network structures as Figure 3.8, except the critical node I , which is impacted by both routes. At this special node, flows from both directions need to be considered. To deal with this, we re-design the maximum flow network, in which, original node I is represented by two nodes I and \tilde{I} (see Figure 3.10), and a new arc (I, \tilde{I}) is added. The capacity of arc (I, \tilde{I}) is constrained by the available green time for intersection (node) I , i.e., $g_{I, i \& j}^a(t)$. At node I , there are two pairs of incoming flows $-\Delta r_{I-1,i}(t)$ & $\Delta g_{I-1,i}(t)$ and $-\Delta r_{(I-1)',j}(t)$ & $\Delta g_{(I-1)',j}(t)$ from node $I-1$ and $(I-1)'$ respectively; there are external flows $d_{I,i}^r(t)$ and $d_{I,j}^r(t)$ for red reductions of two directions and external flows $d_{I,i}^g(t)$ and $d_{I,j}^g(t)$ for green extensions of two directions. The total flow coming out of node I is the summation of flows for both directions, $-\Delta r_{I,i}(t) - \Delta r_{I,j}(t) + \Delta g_{I,i}(t) + \Delta g_{I,j}(t)$, which is constrained by the capacity of arc (I, \tilde{I}) , i.e., $g_{I, i \& j}^a(t)$.

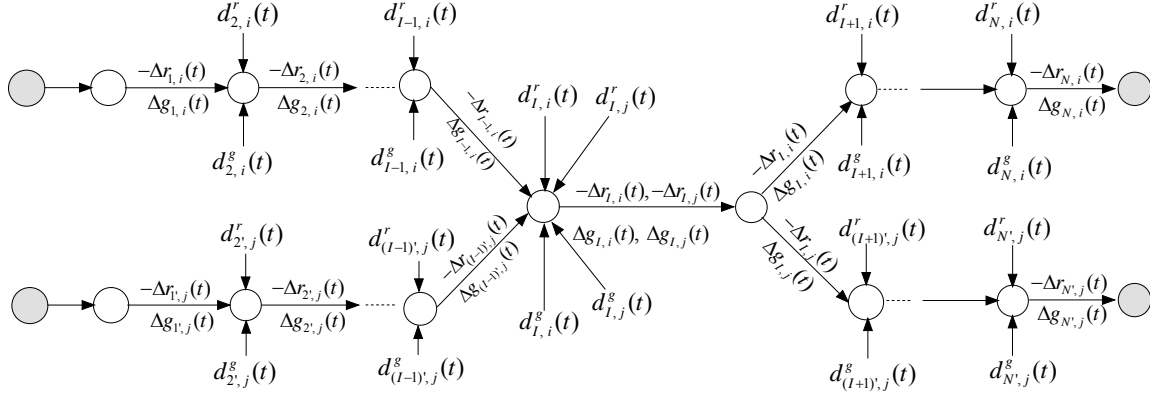


Figure 3.10 Maximum flow network for two intersecting oversaturated routes

3.4 Solution Method – A Forward-Backward Procedure (FBP)

Before introducing the solution method, we first investigate the uniqueness of solution to the proposed maximum flow model Eq. (3.12). Assume $[\Delta r_{1,i}^*(t), \dots, \Delta r_{N,i}^*(t), \Delta g_{1,i}^*(t), \dots, \Delta g_{N,i}^*(t)]$ is the optimal solution of Eq. (3.12), one can verify that $[\Delta r_{1,i}^*(t) + a, \dots, \Delta r_{N,i}^*(t) + a, \Delta g_{1,i}^*(t) + a, \dots, \Delta g_{N,i}^*(t) + a]$, $a \in R$, is also an optimal solution of Eq. (3.12) because it generates the same objective value and satisfies all the constraints as well. It is because that one can shift the offset of every intersection along the oversaturated route by the same amount (a in this case), and that will not change the internal kinematic relations between intersections, such as queue formation and discharging. In other words, a fixed point for intersection offsets is missing for this problem. To make the optimal solution unique, a boundary condition $\Delta r_{1,i}(t) = 0$ is added to the problem Eq. (3.12), indicating that we take the green starting point of the first intersection as the fixed reference point. Similar analysis also applies to the problem Eq. (3.15) for two intersecting oversaturated routes.

To solve the maximum flow based control model, a Forward-Backward Procedure (FBP) is proposed and described in the following. The FBP consists of two processes: a forward process, which is applied along the direction of traffic, and a backward process, which follows the opposite direction.

3.4.1 FBP for One Oversaturated Route

(1) Forward Process (FP)

The forward process aims to eliminate both spillovers and residual queues by reducing red or increasing green of oversaturated phase without considering the constraints from other conflicting phases. The process is applied along the direction of flow and calculates the red and green changes for each oversaturated phase during time period t (i.e. $\Delta r_{n,i}^F(t)$ & $\Delta g_{n,i}^F(t)$). Note that the superscript “ F ” indicates the “Forward” process.

To eliminate spillover, we need to adjust the red time. As discussed, to make the solution unique, the red change for the first intersection is set to zero, i.e. $\Delta r_{1,i}^F(t) = 0$. For any other node, based on the relationship between spillover time and red changes (the first equation in Eq. (3.12)), one can derive the equation $\Delta r_{n,i}^F(t) = \Delta r_{n-1,i}^F(t) - S_{n-1,i}(t-1)$. So the red change for each node can be calculated by the following:

$$\begin{cases} \Delta r_{1,i}^F(t) = 0 \\ \Delta r_{n,i}^F(t) = \Delta r_{n-1,i}^F(t) - S_{n-1,i}(t-1), \quad n = 2, \dots, N \end{cases} \quad \text{Eq. (3.16)}$$

After determining the red time changes, we further adjust green time to eliminate residual queue. In order to find the maximum flow through the network, the flow out of the first node is set to its arc capacity, i.e., $\Delta g_{1,i}^F(t) = g_{1,i}^a(t)$. According to the second constraint in Eq. (3.12), we have the following equation, which generates the green changes for every node.

$$\begin{cases} \Delta g_{1,i}^F(t) = g_{1,i}^a(t) \\ \Delta g_{n,i}^F(t) = \Delta g_{n-1,i}^F(t) + T_{n,i}(t-1) - S_{n-1,i}(t-1) - [g_{n,i}(t-1) - g_{n-1,i}(t-1)], \quad n = 2, \dots, N \end{cases} \quad \text{Eq. (3.17)}$$

Eq. (3.16) and Eq. (3.17) are very intuitive. Eq. (3.16) is to eliminate spillover by changing the duration of red lights. The amount of red time reduction at any intersection should accommodate not only the removal of the spillover to the upstream intersection ($S_{n-1,i}(t-1)$), but also the increase of the arrival flow due to the red reduction made at the upstream intersection ($\Delta r_{n-1,i}^F$). Eq. (3.17) is to extend green time by $T_{n,i}(t-1) - S_{n-1,i}(t-1) - [g_{n,i}(t-1) - g_{n-1,i}(t-1)]$ to discharge residual queues. Similarly, we need to account for the green change of the upstream intersection ($\Delta g_{n-1,i}^F$). The backward process will consider the situation if capacity constraints are violated.

(2) *Backward Process (BP)*

The forward process follows the traffic direction and adds extra green time $\Delta g_{n,i}^F(t) - \Delta r_{n,i}^F(t)$ to discharge the residual queue and to remove spillover for each intersection. However, desired green increases for some intersections may not be achievable due to the other constraints, i.e., green requirement for conflicting phases to discharge queues. To solve this problem, the backward process is designed to gate traffic when the required green time changes calculated in the forward pass are not feasible.

In this process, we start from the last intersection and follow the direction of the opposing flow to check whether the arc flow is less than or equal to arc capacity. The residual capacity $R_{n,i}(t)$ is calculated using Eq. (3.18) at each arc. Positive $R_{n,i}(t)$ means available green can accommodate the required green time increase ($\Delta g_{n,i}^F(t) - \Delta r_{n,i}^F(t)$); negative $R_{n,i}(t)$ means available green is insufficient.

$$R_{n,i}(t) = g_{n,i}^a(t) - (\Delta g_{n,i}^F(t) - \Delta r_{n,i}^F(t)), n = N, \dots, 1 \quad \text{Eq. (3.18)}$$

After calculating the residual capacity for each arc, the backward green time adjustment $\Delta g_i^B(t)$ is equal to the minimum residual capacity for all arcs along the oversaturated route. Since $R_{1,i}(t) = 0$ ($\Delta g_{1,i}^F(t) = g_{1,i}^a(t)$, $\Delta r_{1,i}^F(t) = 0$), the minimum residual capacity is non-positive. If $\min_{n \in \{1, \dots, N\}} [R_{n,i}(t)] = 0$, the requested green time increase ($\Delta g_{n,i}^F(t) - \Delta r_{n,i}^F(t)$) from the forward process will be satisfied at all arcs and no further adjustment is needed in the backward process. However, if $\min_{n \in \{1, \dots, N\}} [R_{n,i}(t)] < 0$, there is at least one arc where the capacity constraint is violated.

The adjustment term $\Delta g_i^B(t)$ ($\Delta g_i^B(t) < 0$) is utilized to make sure the capacity constraints are satisfied at all arcs.

$$\Delta g_i^B(t) = \min_{n \in \{1, \dots, N\}} [R_{n,i}(t)] \quad \text{Eq. (3.19)}$$

The final signal timing changes for every node are calculated by Eq. (3.20). The final red time change is equal to the summation of the calculated value in the forward process $\Delta g_{n,i}^F(t)$ and the adjustment term $\Delta g_i^B(t)$ in the backward process. Note that, $\Delta g_i^B(t)$ is the same for every intersection along the route.

$$\begin{cases} \Delta r_{n,i}(t) = \Delta r_{n,i}^F(t) \\ \Delta g_{n,i}(t) = \Delta g_{n,i}^F(t) + \Delta g_i^B(t) \end{cases}, \quad n \in \{1, \dots, N\} \quad \text{Eq. (3.20)}$$

3.4.2 FBP for a Network with Two Intersecting Oversaturated Routes

With two intersecting oversaturated routes, the available green at intersection I ($g_{I,i\&j}^a(t)$) needs to be split between phase i and j . We can split the available green time proportionally according to the requested green times from the forward process, i.e.,

$$\begin{cases} g_{I,i}^a(t) = g_{I,i\&j}^a(t) \times \frac{g_{I,i}^F(t)}{g_{I,i}^F(t) + g_{I,j}^F(t)} \\ g_{I,j}^a(t) = g_{I,i\&j}^a(t) \times \frac{g_{I,j}^F(t)}{g_{I,i}^F(t) + g_{I,j}^F(t)} \end{cases}$$

However, such split method may not be efficient because the binding constraints for available green times on one or both oversaturated routes may not come from the critical intersection. To overcome such deficiency, we can first compute the residual capacity for all intersections except the critical intersection I and, the backward adjustment term for both directions, $\Delta g_i^B(t)$ and

$\Delta g_j^B(t)$, using equation Eq. (3.19). The “^” sign is used because here we did not consider intersection I . We can then calculate the requested green time increase for phase i and j of intersection I , denoted by $\Delta_{I,i}^R(t)$ and $\Delta_{I,j}^R(t)$.

$$\begin{cases} \Delta_{I,i}^R(t) = \Delta g_{I,i}^F(t) + \Delta g_i^B(t) - \Delta r_{I,i}^F(t) \\ \Delta_{I,j}^R(t) = \Delta g_{I,j}^F(t) + \Delta g_j^B(t) - \Delta r_{I,j}^F(t) \end{cases} \quad \text{Eq. (3.21)}$$

If $g_{I,i\&j}^a(t) \geq \Delta_{I,i}^R(t) + \Delta_{I,j}^R(t)$, then the available green constraint at intersection I is satisfied. The backward process adjustment terms $\Delta g_i^B(t)$ and $\Delta g_j^B(t)$ are equal to $\Delta g_i^B(t)$ and $\Delta g_j^B(t)$ respectively, see Eq. (3.22).

$$\begin{cases} \Delta g_i^B(t) = \Delta g_i^B(t) \\ \Delta g_j^B(t) = \Delta g_j^B(t) \end{cases} \quad \text{Eq. (3.22)}$$

Otherwise, the available green time at intersection I cannot satisfy the total requested green time increase for both directions i and j . The total available green time $g_{I,i\&j}^a(t)$ is split proportionally to two directions, $g_{I,i}^a(t)$ and $g_{I,j}^a(t)$ according to the requested green time increase, see Eq. (3.23). The total backward process adjustment terms for two directions are determined by Eq. (3.24).

$$\begin{cases} g_{I,i}^a(t) = g_{I,i\&j}^a(t) \times \frac{\Delta_{I,i}^R(t)}{\Delta_{I,i}^R(t) + \Delta_{I,j}^R(t)} \\ g_{I,j}^a(t) = g_{I,i\&j}^a(t) \times \frac{\Delta_{I,j}^R(t)}{\Delta_{I,i}^R(t) + \Delta_{I,j}^R(t)} \end{cases} \quad \text{Eq. (3.23)}$$

$$\begin{cases} \Delta g_i^B(t) = g_{I,i}^a(t) - (\Delta g_{I,i}^F(t) - \Delta r_{I,i}^F(t)) \\ \Delta g_j^B(t) = g_{I,j}^a(t) - (\Delta g_{I,j}^F(t) - \Delta r_{I,j}^F(t)) \end{cases} \quad \text{Eq. (3.24)}$$

The final red/green time changes can then be calculated using Eq. (3.20) for both directions.

3.4.3 The Optimality Analysis of the Solution of the FBP

We now prove that the solution generated by the FBP is the optimal solution for the MFP. The proof is divided into two portions: one is for a single oversaturated path and the other is for two intersecting oversaturated paths.

Theorem: The FBP provides optimal solutions to the maximum flow based signal timing control models Eq. (3.12) and Eq. (3.15).

Proof:

(1) For one oversaturated route

Assume $[\Delta g_{1,i}^*(t), \dots, \Delta g_{N,i}^*(t), \Delta r_{1,i}^*(t), \dots, \Delta r_{N,i}^*(t)]$ is the signal control solution provided by FBP.

We first check the feasibility of the solution. Based on Eq. (3.18), Eq. (3.19) and Eq. (3.20),

$$\begin{aligned} \Delta g_{n,i}^*(t) - \Delta r_{n,i}^*(t) &= \Delta g_{n,i}^F(t) + \Delta g_i^B(t) - \Delta r_{n,i}^F(t) \\ &= \Delta g_{n,i}^F(t) - \Delta r_{n,i}^F(t) + \min_{m \in \{1, \dots, N\}} [R_{m,i}(t)] \\ &\leq \Delta g_{n,i}^F(t) - \Delta r_{n,i}^F(t) + R_{n,i}(t) \\ &= g_{n,i}^a(t) \end{aligned}$$

Therefore, the third constraint in Eq. (3.12) is satisfied. Since Eq. (3.16) and Eq. (3.17) are essentially derived from the first two constraints of Eq. (3.12), it is easy to show that

$[\Delta \tilde{g}_{1,i}^*(t), \dots, \Delta \tilde{g}_{N,i}^*(t), \Delta \tilde{r}_{1,i}^*(t), \dots, \Delta \tilde{r}_{N,i}^*(t)]$ also satisfies the first two constraints of Eq. (3.12).

Therefore, the FBP generates a feasible solution to the model Eq. (3.12).

We now show the optimality of the FBP solution by contradiction. Assume that the signal control solution provided by the FBP is not optimal. Therefore, there must exist a control vector

$[\Delta \tilde{g}_{1,i}(t), \dots, \Delta \tilde{g}_{N,i}(t), \Delta \tilde{r}_{1,i}(t), \dots, \Delta \tilde{r}_{N,i}(t)]$ which satisfies all constraints in Eq. (3.12) and

$$\Delta \tilde{g}_{1,i}(t) + \Delta \tilde{r}_{1,i}(t) > \Delta g_{1,i}^*(t) + \Delta r_{1,i}^*(t).$$

Based on the definition of $\Delta g_i^B(t)$ in Eq. (3.19), there exists a node $k \in M$ such that

$$\Delta g_i^B(t) = R_{k,i}(t) = g_{k,i}^a(t) - (\Delta g_{k,i}^F(t) - \Delta r_{k,i}^F(t))$$

which implies $\Delta g_{k,i}^*(t) - \Delta r_{k,i}^*(t) = \Delta g_i^B(t) + (\Delta g_{k,i}^F(t) - \Delta r_{k,i}^F(t)) = g_{k,i}^a(t)$.

Add Eq. (3.16) to Eq. (3.17):

$$\begin{aligned} (\Delta g_{n+1,i}(t) - \Delta r_{n+1,i}(t)) - (\Delta g_{n,i}(t) - \Delta r_{n,i}(t)) &= T_{n+1,i}(t-1) - [g_{n+1,i}(t-1) - g_{n,i}(t-1)], \\ n &\in \{1, \dots, N-1\} \end{aligned}$$

Summing up the above equation from $n=1$ to $n=k-1$, we have:

$$(\Delta g_{k,i}(t) - \Delta r_{k,i}(t)) - (\Delta g_{1,i}(t) - \Delta r_{1,i}(t)) = g_{1,i}(t-1) - g_{k,i}(t-1) + \sum_{m=2}^k T_{m,i}(t-1) := H$$

The above equation implies that

$$\begin{aligned}
\Delta \tilde{g}_{k,i}(t) - \Delta \tilde{r}_{k,i}(t) &= H + (\Delta \tilde{g}_{1,i}(t) - \Delta \tilde{r}_{1,i}(t)) \\
&> H + (\Delta g_{1,i}^*(t) - \Delta r_{1,i}^*(t)) \\
&= \Delta g_{k,i}^*(t) - \Delta r_{k,i}^*(t) \\
&= g_{k,i}^a(t)
\end{aligned}$$

This inequality violates the capacity constraint (available green) of arc k in model Eq. (3.12). Therefore, no solutions can be better than the one provided by FBP.

(2) For two intersecting oversaturated routes

Assume $[\Delta g_{1,i}^*(t), \dots, \Delta g_{N,i}^*(t), \Delta r_{1,i}^*(t), \dots, \Delta r_{N,i}^*(t), \Delta g_{1,j}^*(t), \dots, \Delta g_{N,j}^*(t), \Delta r_{1,j}^*(t), \dots, \Delta r_{N,j}^*(t)]$ is the solution provided by FBP for two intersecting oversaturated paths. Similar approaches can be applied to show the solution satisfies the first three constraints of Eq. (3.15), since it is the same as Eq. (3.12). We just need to show the validity of the final constraint. According to Eq. (3.19),

$$\begin{aligned}
&-\Delta r_{1,i}^*(t) - \Delta r_{1,j}^*(t) + \Delta g_{1,i}^*(t) + \Delta g_{1,j}^*(t) \\
&= -\Delta r_{1,i}^F(t) - \Delta r_{1,j}^F(t) + \Delta g_{1,i}^F(t) + \Delta g_i^B(t) + \Delta g_{1,j}^F(t) + \Delta g_j^B(t)
\end{aligned}$$

If $g_{1,i\&j}^a(t) \geq \Delta_{1,i}^R(t) + \Delta_{1,j}^R(t)$, based on Eq. (3.21) and Eq. (3.22),

$$\begin{aligned}
&-\Delta r_{1,i}^*(t) - \Delta r_{1,j}^*(t) + \Delta g_{1,i}^*(t) + \Delta g_{1,j}^*(t) \\
&= -\Delta r_{1,i}^F(t) + \Delta g_{1,i}^F(t) + \Delta g_i^B(t) - \Delta r_{1,j}^F(t) + \Delta g_{1,j}^F(t) + \Delta g_j^B(t) = \Delta_{1,i}^R(t) + \Delta_{1,j}^R(t) \leq g_{1,i\&j}^a(t) \quad \text{If} \\
&g_{1,i\&j}^a(t) < \Delta_{1,i}^R(t) + \Delta_{1,j}^R(t), \text{ based on Eq. (3.23) and Eq. (3.24),}
\end{aligned}$$

$$\begin{aligned}
&-\Delta r_{1,i}^*(t) - \Delta r_{1,j}^*(t) + \Delta g_{1,i}^*(t) + \Delta g_{1,j}^*(t) \\
&= -\Delta r_{1,i}^F(t) - \Delta r_{1,j}^F(t) + \Delta g_{1,i}^F(t) + \Delta g_i^B(t) + \Delta g_{1,j}^F(t) + \Delta g_j^B(t) = g_{1,i}^a(t) + g_{1,j}^a(t) = g_{1,i\&j}^a(t)
\end{aligned}$$

Therefore, the FBP for two intersecting oversaturated paths provides a feasible solution to the MFP Eq. (3.15).

We now show the optimality under two conditions:

According to Eq. (3.21) and Eq. (3.22),

If $g_{I,i\&j}^a(t) \geq \Delta_{I,i}^R(t) + \Delta_{I,j}^R(t)$, then the capacity constraint at the intersection I is not binding and the maximum flow through the network is determined by other intersections. Similar to the optimality proof for one path, the FBP for two intersecting path generates the optimal solution to the MFP Eq. (3.15).

If $g_{I,i\&j}^a(t) < \Delta_{I,i}^R(t) + \Delta_{I,j}^R(t)$, then the capacity constraint is binding at intersection I . The maximum total flow for two paths is constrained by the total available green time of intersection I . As long as the assigned green time is smaller than the requested green time, i.e.,

$g_{I,i}^a(t) \leq \Delta_{I,i}^R(t)$ and $g_{I,j}^a(t) \leq \Delta_{I,j}^R(t)$, the total flow will be maximum; otherwise, green time will be wasted. According to Eq. (3.23), one can derive

$$\begin{cases} g_{I,i}^a(t) = g_{I,i\&j}^a(t) \times \frac{\Delta_{I,i}^R(t)}{\Delta_{I,i}^R(t) + \Delta_{I,j}^R(t)} = \Delta_{I,i}^R(t) \times \frac{g_{I,i\&j}^a(t)}{\Delta_{I,i}^R(t) + \Delta_{I,j}^R(t)} < \Delta_{I,i}^R(t) \\ g_{I,j}^a(t) = g_{I,i\&j}^a(t) \times \frac{\Delta_{I,j}^R(t)}{\Delta_{I,i}^R(t) + \Delta_{I,j}^R(t)} = \Delta_{I,j}^R(t) \times \frac{g_{I,i\&j}^a(t)}{\Delta_{I,i}^R(t) + \Delta_{I,j}^R(t)} < \Delta_{I,j}^R(t) \end{cases}$$

Therefore, the FBP for two intersecting oversaturated paths generates one optimal solution to MFP Eq. (3.15).

3.5 Simulation Test

A simulation study is conducted to test whether the proposed model can improve the performance of signalized arterials under oversaturated conditions. A network with 5 intersections along the Fair Oaks Ave in the City of Pasadena, CA is built in VISSIM (see Figure 3.11). The length of the corridor is 0.4 mile, the north-south direction is the coordinated direction and the speed limit is 30 MPH. In order to test the performance of the proposed model, we create a southbound flow surge in the middle period of the simulation. The normal flow condition is shown in Figure 3.11 (a) with a southbound flow rate 1500 VPH and the increased flow condition is shown in Figure 3.11 (b) with a southbound flow rate 3000 VPH. As demonstrated by Table 3.1, during the first half an hour of the simulation (0~1800 Seconds), the flow condition is normal; the southbound flow rate is increased from 1500 VPH to 3000 VPH in the middle one hour (1800~5400 Seconds); finally, the network input flows get back to normal in the final half an hour (5400~7200 Seconds). The test scenario is designed to represent the traffic condition when there is a demand surge due to some unexpected reasons.

Synchro was first utilized to optimize the signal timings according to the normal and increased traffic flow conditions shown in Figure 3.11. With the normal flow, the optimized cycle length is 80 seconds; with the increased southbound flow, the optimized cycle length is 120 seconds. According to this, four types of control strategies are designed (see Table 3.2). Control strategy

#1 and #2 represent the optimal signal timings according to normal flow and increased flow conditions respectively. The two strategies are both actuated-coordinated and are implemented through the Ring-Barrier-Controller (RBC) in Vissim. Control strategy #3 and #4 follow the Forward-Backward Procedure with different cycle lengths (80 and 120 seconds) when oversaturated conditions are detected along the corridor. The control interval is set to 2 cycles. Since the signal timings need to be changed in real time for the FBP control to respond to the latest performance measures, i.e., TOSI and SOSI values, control strategy #3 and #4 are implemented through the COM interface in Vissim. When there is no oversaturation detected, control strategy #3 and #4 basically implement fixed-timing control with the optimal parameters from Synchro.

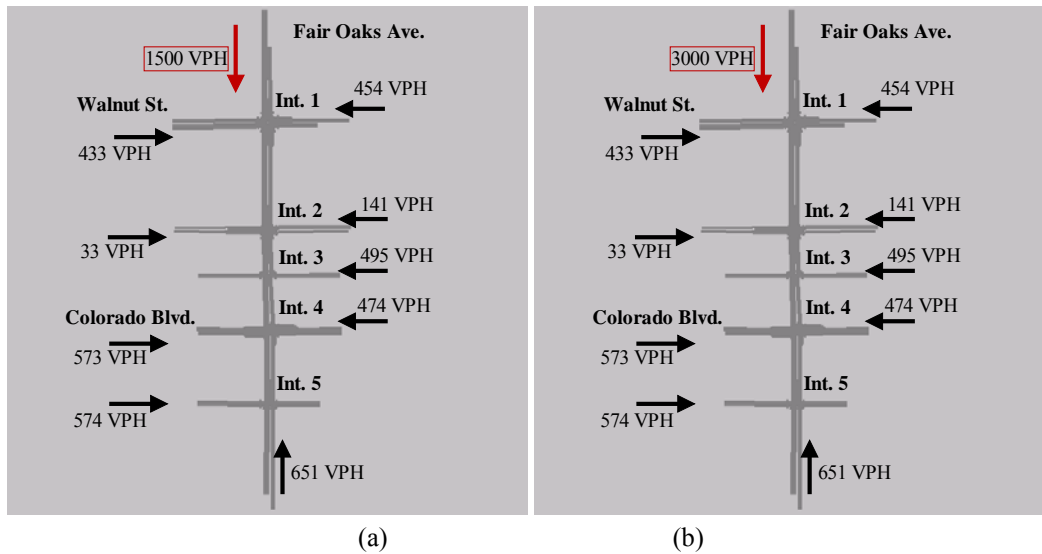


Figure 3.11 Vissim simulation network (a) normal flow condition (b) increased flow condition

Table 3.1. Traffic flow conditions during the simulation period

Simulation time (Sec)	Traffic Flow Conditions
0~1800	Normal flow condition (a)
1800~5400	Increased flow condition (b)
5400~7200	Normal flow condition (a)

Table 3.2. Control strategy comparison

Control Strategy No.	Description	Cycle Length (Sec)
1*	Actuated-coordinated control	80
2	Actuated-coordinated control	120
3*	FBP	80
4	FBP	120

To compare the performance of different control strategies, the simulation was run 5 times with 5 different random seeds under each control strategy and the average performance of each control

strategy over different random seeds is summarized below. One should note that, in real field, since the real-time demand information is very difficult to get or be estimated, it is almost impossible to change the signal timings to the corresponding optimal settings when demand changes. It is more likely and more realistic that when demand surge like Figure 3.11 (b) happens, the implemented control strategy in field will still be #1 (optimized according to the normal flow conditions). Therefore, it makes more sense to compare the performance of control strategy #1 and #3. However, for research purpose all the results under different control strategies are listed below and the performance of strategy #1 is considered as the base line.

Table 3.3 and Figure 3.12 summarize the network performance during the whole simulation period under different strategies. For the average delay per vehicle, strategy #1 ends up with 81.37 seconds. With the optimal signal timings according to increased flow, strategy #2 reduces the number to 61.73 seconds, which is a 24.14% decrease. The FBP with a cycle length of 80 seconds reduces the average delay to 64.28 seconds (21% decrease), while the FBP with a cycle length of 120 seconds reduces the average delay to 56.95 seconds (30.02% decrease). For the average number of stops and average speed, similar trend can also be found where strategy #2, #3 and #4 substantially outperform strategy #1.

Table 3.3. Network performance comparison

	Strategy #1	Strategy #2		Strategy #3		Strategy #4	
		Value	(%)	Value	(%)	Value	(%)
Average Delay (Seconds/per veh.)	81.37	61.73	-24.14	64.28	-21.00	56.95	-30.02
Average # of stops (per veh.)	2.05	1.43	-30.34	1.60	-21.96	1.25	-39.12
Average Speed (MPH)	10.95	13.19	+20.42	12.92	+17.96	13.77	+25.76

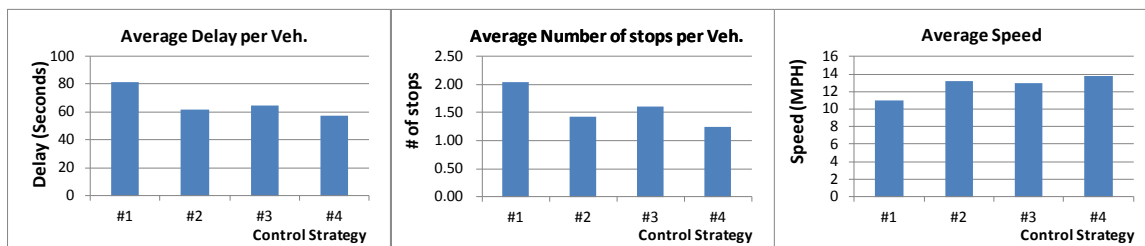


Figure 3.12 Comparison of network performance

Figure 3.13 compares the trajectories of the southbound throughputs under different control strategies. Each point represents the total throughputs during a 5-minute interval. One can see that, in the middle one hour period when the southbound input flow has a large increase, strategy #2, #3 and #4 have much better performances comparing with strategy #1 and within each interval more vehicles can be discharged through the southbound exit under strategy #2, #3 and

#4. The total throughputs of different exits of the network in the 2-hour simulation period are summarized in Table 3.4. Seven groups of exits are listed here, where the first group represents the throughput of the southbound exit, the second indicates the throughput of northbound exit and the remaining five groups represent the respective throughputs of the side streets at each intersection along the route. One can see that the throughput of the southbound exit for the two-hour period is 3021.2 vehicles under strategy #1. However, with strategy #2, #3 and #4 respectively, the total throughput is increased to 3716.83 (23.03% increase), 3762.0 (24.52% increase) and 3809.8 (26.10% increase). Overall, when comparing with the strategy #1, the total throughput of the whole network is increased by 9.0% under strategy #2, 9.6% under strategy #3 and 10.16% under strategy #4.

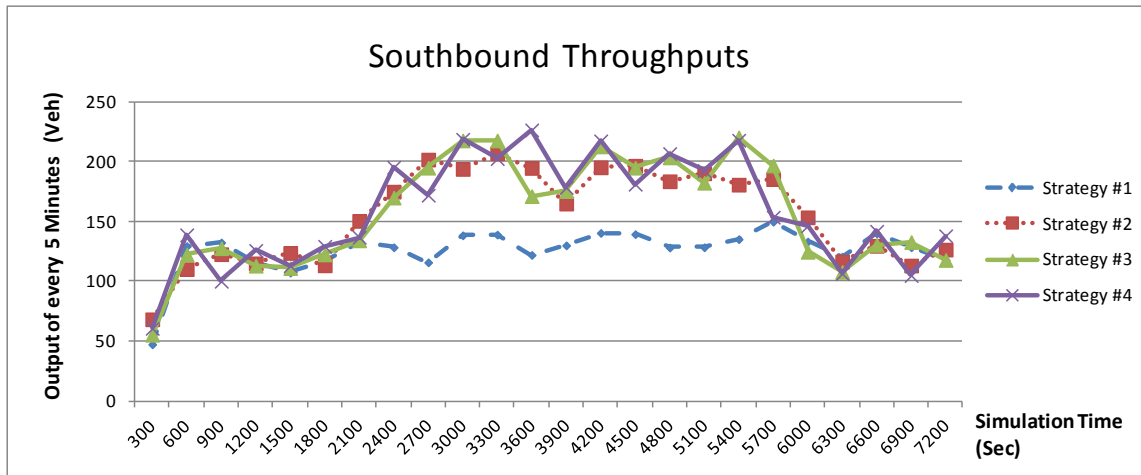


Figure 3.13 Comparison of southbound throughputs

Table 3.4. Two-hour throughputs comparison

	Strategy #1	Strategy #2		Strategy #3		Strategy #4	
		Value	(%)	Value	(%)	Value	(%)
Southbound	3021.20	3716.83	+23.03	3762.00	+24.52	3809.80	+26.10
Northbound	1248.80	1244.83	- 0.32	1242.60	- 0.50	1246.60	- 0.18
Int. 1 side streets	1490.80	1546.67	+3.75	1539.60	+3.27	1544.80	+3.62
Int. 2 side streets	647.40	762.17	+17.73	772.20	+19.28	776.40	+19.93
Int. 3 side street	1120.60	1180.67	+5.36	1180.20	+5.32	1187.00	+5.93
Int. 4 side streets	1795.40	1800.50	+0.28	1815.40	+1.11	1806.00	+0.59
Int. 5 side street	1555.80	1609.33	+3.44	1613.80	+3.73	1614.60	+3.78
TOTAL	10880.0	11861.0	+9.0	11925.8	+9.6	11985.2	+10.16

Figure 3.14 compares the side streets' maximum queue length at two bottleneck intersections where the side street input demand is comparatively higher. One can see that, generally shorter cycle length generates shorter maximum queue length in each cycle. For example, strategy #1 has shorter maximum queue than strategy #2 and strategy #3 has shorter maximum queue than strategy #4. The side streets' maximum queue length in each cycle under strategy #1 and #2 has less fluctuation than the one under strategy #3 and #4. It is because that strategy #1 and #2 has the same control parameters through the whole simulation period, however, when the demand surge happens, strategy #3 and #4 will try to increase the discharging capacity of the oversaturated route by cutting some green time from side streets which may significantly increase the maximum queue length on side streets. But one can see from Figure 3.14, these increases are still within the acceptable ranges.

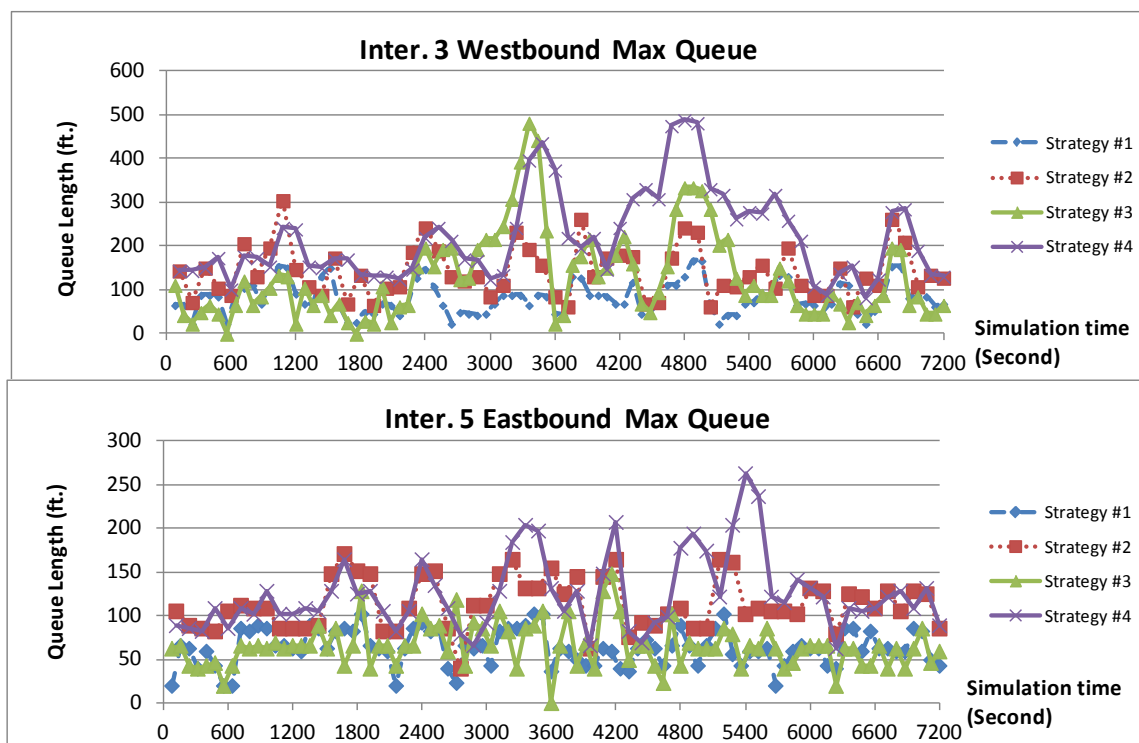


Figure 3.14 Comparison of side streets' maximum queue length in each cycle

To summarize, when oversaturation happens due to demand surge, the strategy #2, #3 and #4 perform much better than the strategy #1 in terms of average delay per vehicle, average number of stops per vehicle, average speed and total throughputs. As discussed before, in real-field applications, real-time demand information is almost impossible to measure or estimate, especially when the demand is changing dramatically along time. However, the proposed FBP can adjust signal timings based on the measured Oversaturation Severity Index and does not rely on the demand information. Therefore, it makes more sense to compare the performance of strategy #1 and #3. As one can see from the results above, the proposed FBP outperform Synchro in handling oversaturated traffic conditions, because the FBP systematically considers the discharging capacities between intersections and side street constraints along the oversaturated

route. By applying the FBP, oversaturated conditions between intersections can be alleviated to the most extent, the traffic along the oversaturated route becomes much smoother and the total throughput can be significantly improved.

Chapter 4. Develop Integrated Control Strategies

4.1 Problem Statement

This report aims to solve the diversion control problem between two alternative routes in order to fully utilize the available capacities. A more typical and challenging situation is that the two routes belongs to different control types, e.g. one route is freeway and the other is signalized arterial, as shown in Figure 4.1. The two origins O_1 and O_2 might be the same or different and so do the two destinations D_1 and D_2 . In practice, most daily commuters would like to choose one of the routes based on their driving experience and preference. However, if the performance on one of the routes is significantly worse than the other, which might be caused by either recurrent (e.g. daily congestion during peak hours) or non-recurrent (e.g. car crash) event, to divert a portion of travelers to the alternative route with better performance would certainly benefit the whole system. Considering the diversion control between freeway and signalized arterial, when freeway congestion occurs, the control strategy is to divert the freeway traffic to the arterial system. How to inform travelers with real-time traffic information and how to predict the potential impacts of diverting traffic to the diverting route are the two most important questions which need to be answered here.

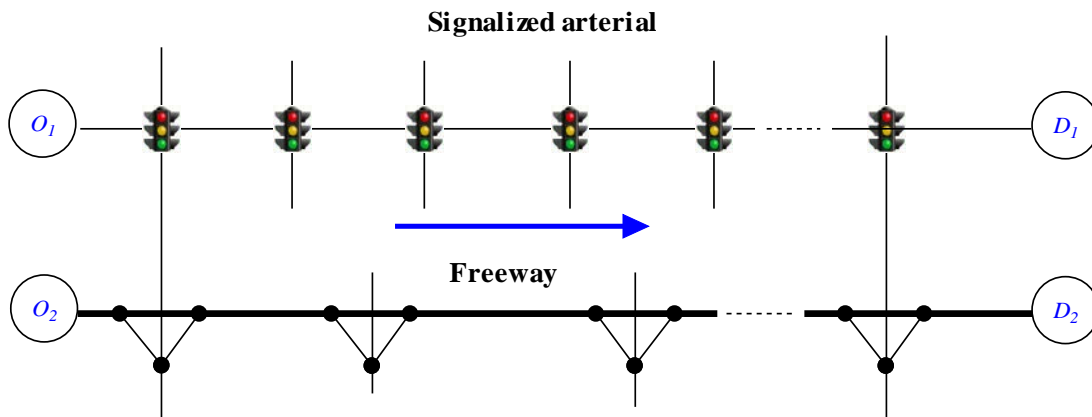


Figure 4.1 Problem statement

4.2 Model Formulation

4.2.1 Performance Estimation

In order to make correct control decisions, the performance of both routes needs to be monitored in real-time. At the end of each control period t , control decisions for the next control period $t+1$ will be made based on the traffic conditions in the immediate past control period t . The control interval usually includes 2~3 signal cycles. In this sub-section, the performance estimation method for both freeway and signalized arterial will be introduced.

(1) Freeway performance estimation

Density and travel time are the two most important measures to reflect freeway performance. To estimate the real-time density and travel time on freeway, certain detection system (e.g. loop detectors, cameras, blue tooth technology and etc) is assumed to be available. Loop detector is one of the most commonly used techniques in the current traffic infrastructure. Detector stations are usually placed every 0.5 to 1 mile along freeways. The loop detector data, such as volume and occupancy, is transferred back to the control center in aggregated levels (e.g. every 30 seconds). In the proposed control model, a freeway corridor is divided into segments such that each segment contains one detector station. The performance of each segment is estimated based on the collected data from corresponding detector station.

Assume the freeway is divided into M segments, $1, 2, \dots, M$. The density of each segment, denoted by $k_m(t)$ (Vehicles/Mile), can be calculated by Eq. (4.1), where $\theta_m(t)$ ($\theta_m(t) \in [0, 1]$) is the average occupancy during control period t , L_v is the average vehicle length (in feet) and L_d is the length of detector (in feet). Then, the average speed of each segment, denoted by $v_m(t)$ (Miles/Hour), can be generated by Eq. (4.2), where $q_m(t)$ is the average 30-second volume during control period t .

$$k_m(t) = \frac{5280 \times \theta_m(t)}{L_v + L_d}, m \in \{1, 2, \dots, M\} \quad \text{Eq. (4.1)}$$

$$v_m(t) = \frac{120 \times q_m(t)}{k_m(t)}, m \in \{1, 2, \dots, M\} \quad \text{Eq. (4.2)}$$

Thus, the travel time along the freeway corridor, denoted by $T^f(t)$, can be calculated by Eq. (4.3), where l_m^f is the length for segment m .

$$T^f(t) = \sum_{m=1,2,\dots,M} [l_m^f / v_m(t)] \quad \text{Eq. (4.3)}$$

(2) Arterial performance estimation

In order to estimate the arterial performance in real-time, the arterial traffic data collection system is also expected to be available, for instance, the SMART-SIGNAL system (Liu, et al., 2009), which automatically archives the event-based high-resolution traffic data (i.e. signal changes and vehicle actuations). Based on the collected data set, real-time second-by-second queue length can be estimated with very high accuracy.

Assume the arterial has N signalized intersections and the queue length (in ft.) for the diverting traffic direction (i.e. phase i) at intersection n at any given second τ is $Q_{n,i}(\tau)$. It can be

estimated using the method introduced in *Liu, et al., 2009*. For any cycle, the average delay $d_{n,i}$ (in Second/Vehicle) can be calculated by Eq. (4.4).

$$d_{n,i} = \frac{1}{A_{n,i} \times h} \sum_{\tau=1}^c \max [Q_{n,i}(\tau) - W_{n,i}(\tau), 0] \quad \text{Eq. (4.4)}$$

Where h is the average space headway of vehicles in queue, c is the cycle length and $A_{n,i}$ is the arrival vehicles for phase i during the cycle. $W_{n,i}(\tau)$ is the location of discharging wave at any given second τ . $W_{n,i}(\tau)$ is calculated by Eq. (4.5), where $\omega_{n,i}$ is the discharging wave speed (in ft./s) and $G_{n,i}^s$ is the green start time.

$$W_{n,i}(\tau) = \begin{cases} \omega_{n,i} \times (\tau - G_{n,i}^s) & \text{if } \tau > G_{n,i}^s \\ 0 & \text{else} \end{cases} \quad \text{Eq. (4.5)}$$

During control period t , the average delay, denoted by $d_{n,i}(t)$, is the average of $d_{n,i}$ over all cycles. Thus, the average travel time for signalized arterial during control period t is

$$T^a(t) = \sum_{n=1}^N \left(\frac{l_{n,i}^a}{v_{n,i}^a} + d_{n,i}(t) \right) \quad \text{Eq. (4.6)}$$

Where $l_{n,i}^a$ is the link length of approach i at intersection n and $v_{n,i}^a$ is the free flow speed.

Before making any adjustment, the residual capacity of each intersection needs to be calculated. Assume that there are N intersections along the signalized arterial, the residual capacity of intersection n (denoted by $\eta_{n,i}^a(t)$) for the phase of diverting traffic direction (i.e. phase i) during the control period t can be calculated by Eq. (4.7). $g_{n,i}(t)$ is the green time for phase i of intersection n during control period t , $s_{n,i}$ is the corresponding saturation flow rate, $z_{n,i}$ is number of lanes and $\gamma_{n,i}(t)$ is the average cycle discharging volume for phase i of intersection n during control period t . The residual capacity measures how much more traffic can be discharged during one cycle at specific intersection based on the current traffic condition.

$$\eta_{n,i}^a(t) = g_{n,i}(t) s_{n,i} z_{n,i} - \gamma_{n,i}(t), n \in \{1, 2, \dots, N\} \quad \text{Eq. (4.7)}$$

The residual capacity along the signalized arterial is the minimum residual capacity among all intersections,

$$\eta^a(t) = \min_{n \in \{1, 2, \dots, N\}} [\eta_{n,i}^a(t)] \quad \text{Eq. (4.8)}$$

When traffic along signalized arterial becomes congested, oversaturated traffic conditions may happen, which will cause detrimental effects to signal operation. An Oversaturation Severity Index (OSI) was proposed by *Wu et al. (2010)* to quantify the severity level of oversaturation by measuring its detrimental effects. Detrimental effect is characterized by either a residual queue at the end of a cycle or a spillover from downstream traffic, both of which create “unusable” green time. In the case of residual queue, the “unusable” green time is the equivalent green time to discharge the residual queue in the following cycle, but for spillover, the “unusable” green time is the time period during which an downstream link is blocked therefore the discharge rate is zero. OSI is further differentiated into *TOSI* (Temporal Oversaturation Severity Index, caused by the residual queue that creates the detrimental effect in temporal dimension) and *SOSI* (Spatial Oversaturation Severity Index, caused by the spillover that creates the detrimental effect in spatial dimension). In the following, we use $S_{n,i}(t)$ to represent the unusable green time caused by spillover at intersection n phase i during time period of t , and use $T_{n,i}(t)$ to represent the unusable green time caused by residual queue.

4.2.2 Diversion Control

Diversion control decisions are made based on the real-time estimated performance of both routes. When the performance on freeway is worse than that on arterial at certain control period t , we may want to divert some traffic from freeway to arterial, see Figure 4.2. This condition can be expressed by Eq. (4.9), where $T^{f \rightarrow a}(t)$ is the diversion cost from freeway to arterial, i.e. travel time on diverting links. In this case, a variable message sign (VMS) can be shown on the freeway side before the diverting point, indicating the travel times on both routes and advising drivers to use the arterial system.

$$T^f(t) > T^{f \rightarrow a}(t) + T^a(t) \quad \text{Eq. (4.9)}$$

Since there is usually no signal control on the freeway mainline, the exact number of diverting traffic is difficult to control. To overcome this problem, we assume drivers are rational and their perception errors follow a standard Gumbel distribution. Then, a Logit decision model is used to predict the diversion rate $\varphi(t+1)$ at the next control period $t+1$, i.e. the percentage of vehicles that will be diverted from freeway to arterial because of the provision of traffic information of both routes. In the model, the diversion rate $\varphi(t+1)$ is calculated based on travel time difference between arterial and freeway as shown in Eq. (4.10), where u is the travel time difference in minutes, β is the coefficient that values travel time with respect to travel utility, and α is the parameter that represents every other factor not related to time, such as drivers’ inertia of mind (i.e. unwillingness to divert). Both α and β can be estimated based on historical data and experience.

$$\varphi(t+1) = \frac{1}{1 + e^{\alpha + \beta u(t)}} \quad \text{Eq. (4.10)}$$

$$u(t) = T^f(t) - [T^a(t) + T^{f \rightarrow a}(t)]$$

The diverted traffic volume from freeway to arterial in the next control period $t+1$, denoted by $\pi^{f \rightarrow a}(t+1)$, can be predicted by Eq. (4.11), where Δt is the control interval. The prediction is based on the assumption that the incoming traffic during control period $t+1$ at the freeway is the same as that during control period t , denoted by $A^f(t)$. Because of the diversion traffic into the arterial system, the signal timings along arterial may need to be adjusted accordingly.

$$\pi^{f \rightarrow a}(t+1) = A^f(t)\phi(t+1)\Delta t \quad \text{Eq. (4.11)}$$

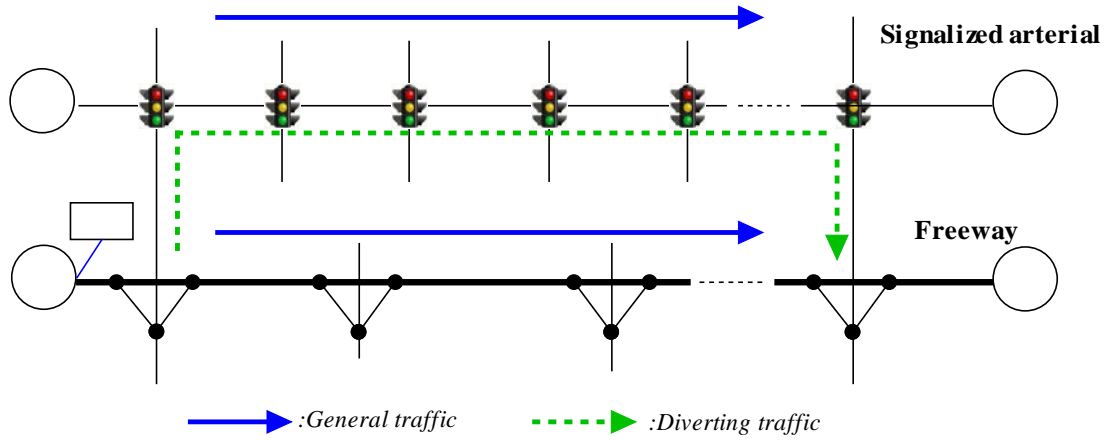


Figure 4.2 Diversion control from freeway to arterial

If $c \times \pi^{f \rightarrow a}(t+1) / \Delta t \leq \eta^a(t)$, where c is the cycle length, the diverting traffic can be handled by the current signal timings along the arterial; however, if $c \times \pi^{f \rightarrow a}(t+1) / \Delta t > \eta^a(t)$, the diverting traffic will cause residual queue at some intersection(s). The residual queue $\psi_{n,i}(t+1)$ at each intersection during the next control period $t+1$ can be predicted by Eq. (4.12), where $\Delta\lambda_{n,i}(t+1)$ is the predicted increase of arrival traffic at intersection n during control period $t+1$. The initial condition is $\Delta\lambda_{1,i}(t+1) = c \times \pi^{f \rightarrow a}(t+1) / \Delta t$. The first equation basically says if the increase of arrival flow at specific intersection during control period $t+1$ (i.e. $\Delta\lambda_{n,i}(t+1)$) is larger than the corresponding residual capacity (i.e., $\eta_{n,i}^a(t)$), there will be a residual queue $\Delta\lambda_{n,i}(t+1) - \eta_{n,i}^a(t)$; otherwise, there will be no residual queue at intersection n . The second equation updates the increase of arrival flow to the downstream intersection (i.e. $\Delta\lambda_{n+1,i}(t+1)$), which is equal to the minimum of the residual capacity (i.e., $\eta_{n+1,i}^a(t)$) and the increase of arrival flow (i.e. $\Delta\lambda_{n,i}(t+1)$) at the current intersection.

$$\begin{cases} \psi_{n,i}(t+1) = \max\left[0, \Delta\lambda_{n,i}(t+1) - \eta_{n,i}^a(t)\right] \\ \Delta\lambda_{n+1,i}(t+1) = \min\left[\eta_{n,i}^a(t), \Delta\lambda_{n,i}(t+1)\right] \end{cases}, n \in \{1, 2, \dots, N\} \quad \text{Eq. (4.12)}$$

When residual queue happens at signalized intersections, it means the current discharging capacity cannot accommodate the increase of traffic. If the signal timings are not properly adjusted, more severe oversaturated conditions, such as spillovers, will appear. Therefore, the maximum flow based signal control model, which is introduced in Chapter 3, is utilized to mitigate or eliminate oversaturated traffic conditions between intersections.

4.3 Simulation Test

In order to test the proposed approach, a case study site was selected in Minneapolis, MN. The site has two major routes (see Figure 4.3), i.e. Trunk Highway 55 (a coordinated high speed signalized arterial) and Interstate freeway 394, connecting the west suburban living areas and the downtown Minneapolis. The total length of the corridor is about 3.5 miles and both routes (i.e. I-394 and TH 55) have a speed limit of 55 MPH. The coordination of the TH 55 favors the eastbound traffic during the AM peak hours because of the large traffic from home to work and it favors the westbound during the PM peak hours to handle the returning traffic. Based on the detector station locations in the field, the I-394 freeway is divided into 6 segments (see Figure 4.3) such that each segment contains one detector station. Figure 4.4 shows the flow-density diagram from the detector station at segment 4 based on the field collected data (the 30-second freeway data) between 6/15/2009 and 6/19/2009. One can easily find out that the critical density for segment 4 is about 150 Vehicles/Mile.

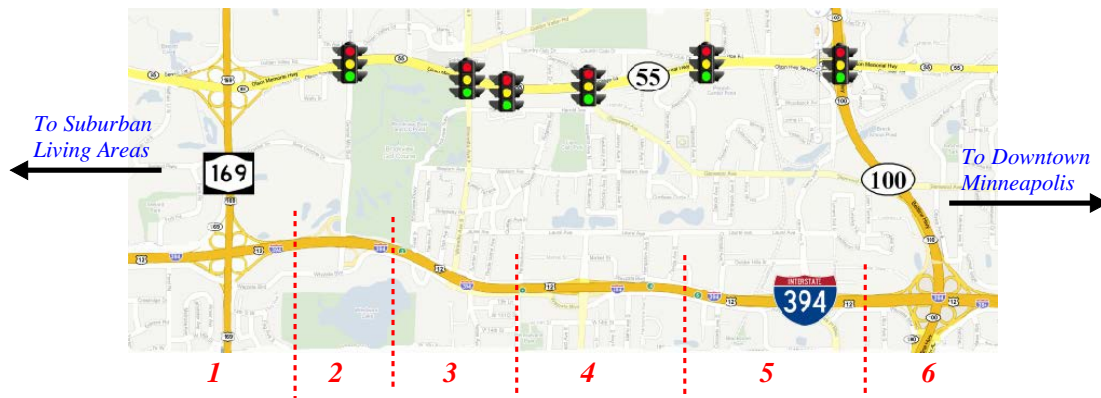


Figure 4.3 Case study site: the TH 55/I-394 corridor, Minneapolis, MN (Source: Google Maps)

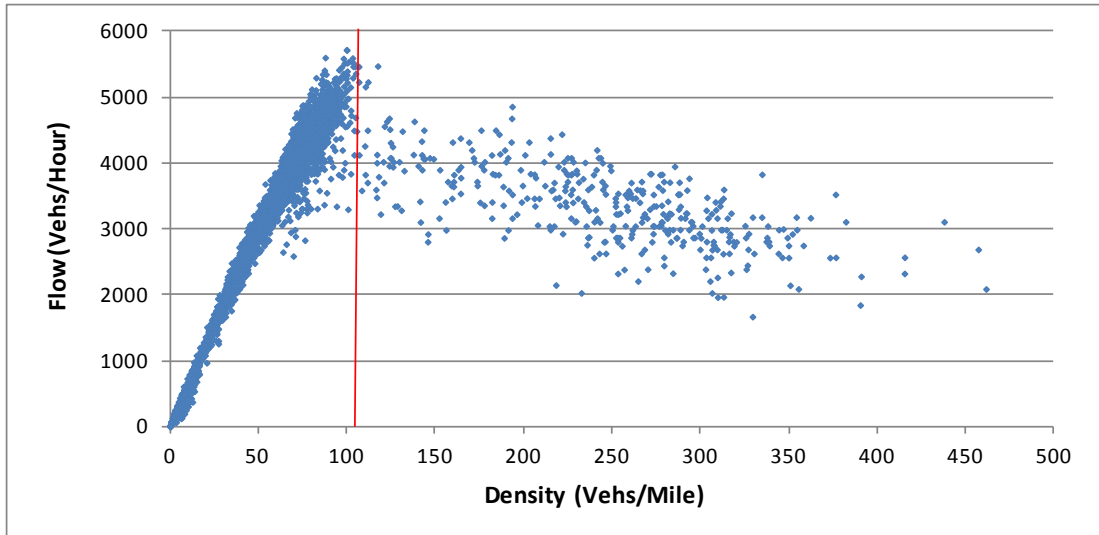


Figure 4.4 Flow-density diagram from three detectors at segment 4

A VISSIM model is then built and calibrated using the field data collected during the morning peak hours (7:00 AM – 9:00 AM) between 6/15/2009 and 6/19/2009, see Figure 4.5. The diverting route is shown in Figure 4.5 by the green dotted line, which goes through TH 169 northbound, TH 55 eastbound and then TH 100 southbound. The simulation control program was written in C# and it controls the simulation in real-time through the COM interface of VISSIM. At each control period, the travel time on freeway (i.e. $T^f(t)$) is estimated based on the segment speed; the travel time on the arterial (i.e. $T^a(t)$) is estimated through the virtual probe approach; the diverting cost $T^{f \rightarrow a}(t)$ is the summation of travel times on TH 169 northbound and TH 100 southbound, which can also be estimated through the same approach as freeway.

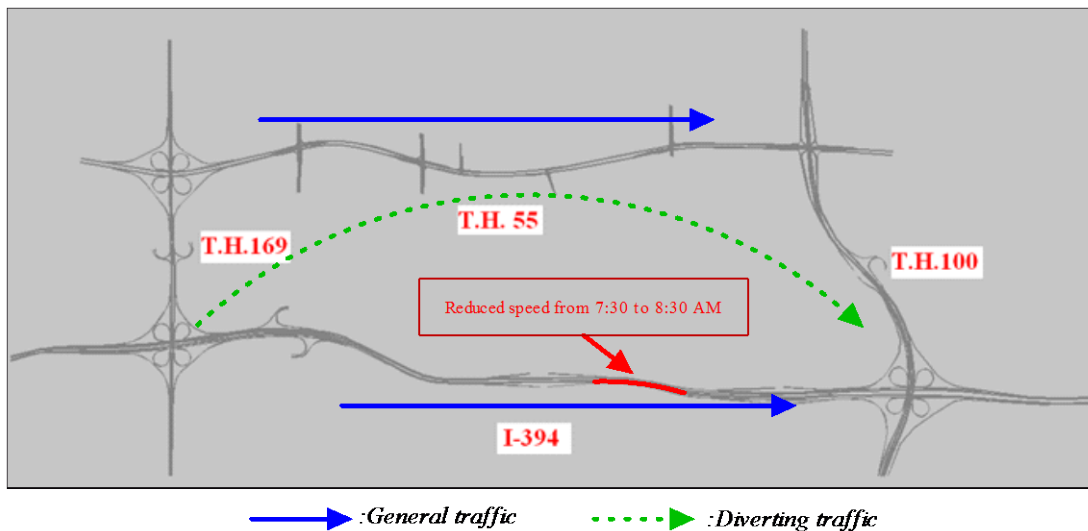


Figure 4.5 VISSIM network of the TH 55/I-394 corridor

The simulation lasts for two hours (7:00 AM – 9:00 AM) and Figure 4.6 shows the demand profiles of the major directions (i.e. I-394 EB, I-394 WB, TH 55 EB and TH 55 WB) for the whole simulation period. The demand profile on I-394 is estimated based on the 30-second freeway data set and the demand profile on TH 55 is generated by the SMART-Signal system. The cycle length of the signalized arterial is 180 seconds and the control interval Δt is 360 seconds. To simulate some unexpected incident (i.e. car crash) happening on freeway, a reduced speed area (10 MPH) with a length of 800 ft is created on the eastbound of I-394 from 7:30 AM to 8:30 AM (see Figure 4.6). Vehicles passing that area during that time window have to reduce their speed and as a result severe congestion will happen on the eastbound of I-394.

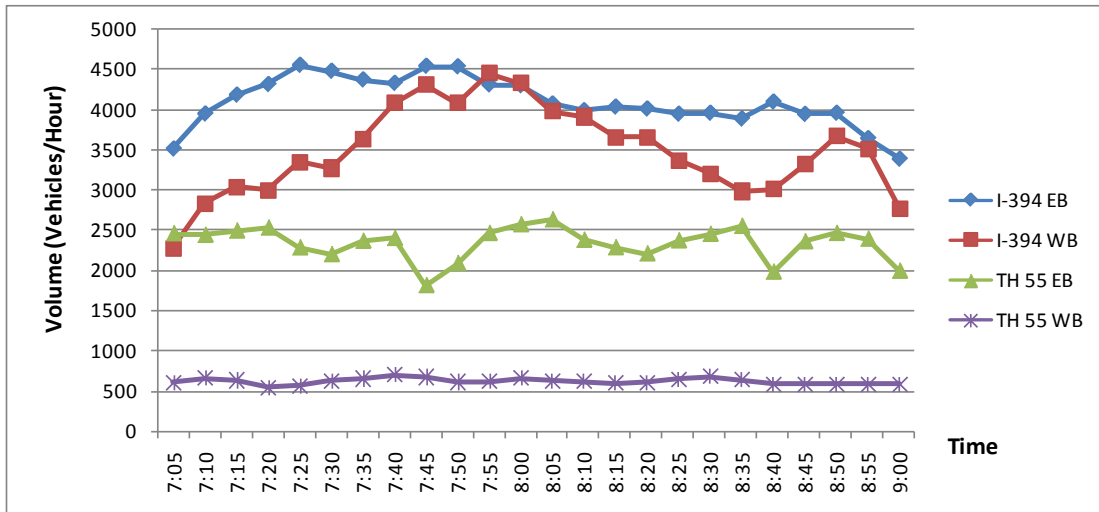


Figure 4.6 Demand profiles for the simulation period

In the following, two scenarios will be tested: one is the base scenario with original control strategy (i.e. independent control) and the other is the scenario with the proposed integrated control strategy. Each scenario is run for 10 times using different random seeds and the average results are listed below. Figure 4.7 shows the travel time profiles of the general route and diverting route. Under the base scenario with the original control strategy, the travel time of the general route (see the blue dashed line with diamond markers) increase dramatically after 7:30 AM when the congestion on freeway happens. The travel time of the diverting route (see the red dashed line with square markers) is relatively consistent during the whole period. Under the integrated control strategy, although the travel time of the general route (see the green line with triangle markers) still increases largely after 7:30 AM, the increasing trend is much slower because of diversion control; on the other hand, the diverting traffic makes the travel time of the diverting route (see the purple solid line with star markers) much higher comparing with that of base scenario.

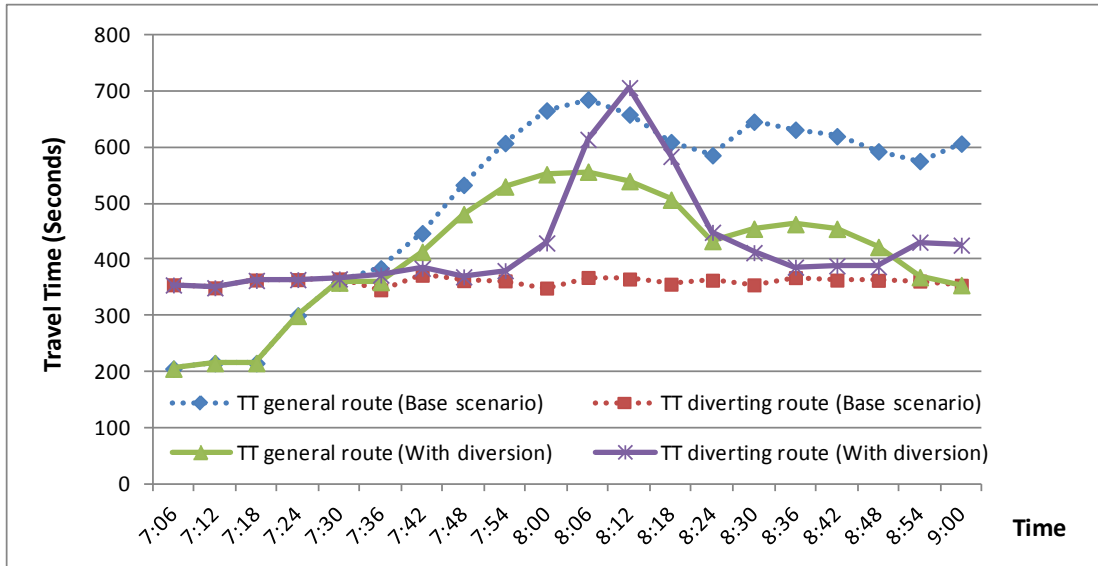


Figure 4.7 Travel times of general route and diverting route under different scenarios

Figure 4.8 presents the relationship between the travel times on different routes and the diversion rate. One can see that, at the beginning of the simulation, the travel time on the general route is lower than that on the diverting route, so no diversion control is needed; after the severe congestion happens on freeway, the travel time on the general route becomes higher than that on the diverting route, a portion of the traffic decides to use the diverting route, which will inevitably increase the travel time on the diverting route. The diversion control continues until the traffic situation reaches the point when the travel time on the diverting route becomes higher than that of the general route. And then, this process repeats. It is not difficult to find that the two lines (travel time on the general route and the diverting route) will weave with each other to best utilize the corridor capacities.

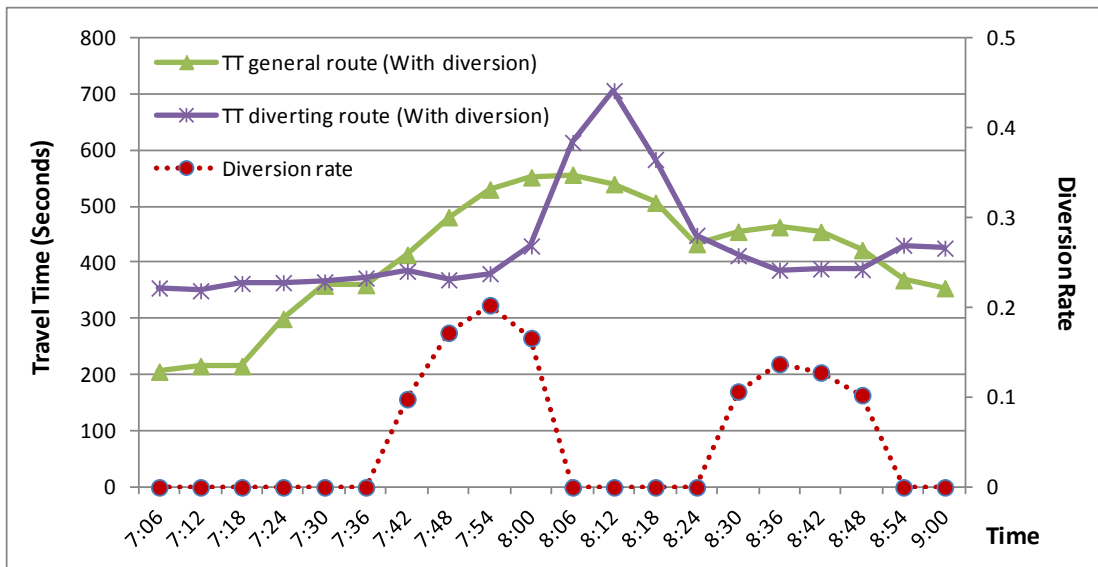


Figure 4.8 Travel time and diversion rate

Table 3.3 summarizes the network performance during the whole simulation period. With original control strategy, the average delay is 55.69 Seconds/Veh, while with the proposed integrated control strategy, the average delay is reduced to 41.14 Seconds/Veh, which is a 26.13% reduction. For average number of stops of the whole network, the proposed control model reduces it from 2.21 to 1.28, a 42.13% reduction. The average speed is increased from 42.12 MPH to 45.86 MPH.

Table 4.1. Network performance comparison

	Base Scenario	With diversion	Change
Average Delay (Seconds per veh.)	55.69	41.14	-26.13%
Average # of stops (per veh.)	2.21	1.28	-42.13%
Average Speed (MPH)	42.12	45.86	+8.89%

In order to test the performance of the proposed control strategy to handle different demand levels, we increase and decrease the mainline demand (i.e. the demand shown in Figure 4.6) by 5% and then run the simulation again.

Table 4.2 presents the network performance under demand variations. When the mainline demand is increased by 5%, the whole network becomes more congested, which can be reflected by the increase of average delay and average number of stops and the decrease of average speed. However, with the proposed diversion control strategy, average delay and average number of stops can be reduced by 16.31% and 38.2% respectively and average speed can be increased by 7%. On the other hand, when the mainline demand is decreased by 5%, the proposed diversion control strategy can still significantly improve the network performance, i.e. reduce average delay by 29.67%, reduce average number of stops by 47.97% and increase average speed by 7.82%. Based on the results discussed above, one can see that the proposed integrated control model can effectively reduce network congestion and smooth traffic movement by utilizing the available capacity along parallel route.

Table 4.2. Network performance comparison with demand variations

	Increase demand by 5%			Decrease demand by 5%		
	Base Scenario	With diversion	Change	Base Scenario	With diversion	Change
Average Delay (Seconds per veh.)	76.79	64.27	-16.31%	40.27	28.32	-29.67%
Average # of stops (per veh.)	3.45	2.13	-38.20%	1.41	0.73	-47.97%
Average Speed (MPH)	37.63	40.26	+7.00%	45.98	49.58	+7.82%

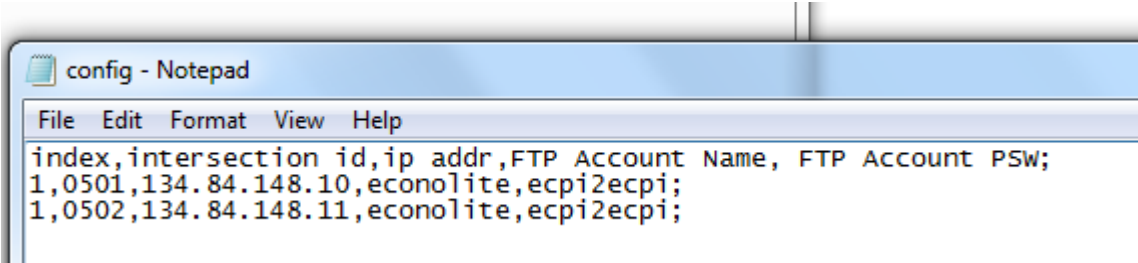
Chapter 5. Installation of the SMART Signal System on TH 55

The new version of the SMART-Signal data collection software can directly retrieve event-based traffic data from the Econolite ASC/3 controllers without additional hardware instrument. This chapter describes the design details of the software and the implementation of the SMART-Signal system in the 10 intersections along TH 55, Minneapolis, MN.

5.1 The New Data Collection Software

The Econolite ASC/3 controllers have a data-logging feature which can store the event-based traffic data into a circular buffer. Once it is enabled, the controller will continuously archive the event-based traffic data including detector actuations, traffic signal changes as well as facility diagnosis information. A built-in FTP server is running on the controller which can be accessed to retrieve the data. The new version of the SMART-Signal data collection software is developed based on such feature of the Econolite ASC/3 controllers. It is deployed in the central server in order to retrieve data from field intersections.

At the initial stage, the software will read a configuration file that contains a list of intersection IDs, controller IP addresses, FTP account names and passwords. An example of the configuration file can be found in Figure 5.1. By default, the account name and password for the built-in FTP server are "econolite" and "ecpi2ecpi" respectively. The hourly log files (with an extension of ".dat") are located in the folder of "/set1" under the root folder. And then, the data collection software will enter an infinite loop to retrieve the event-based traffic data from each controller, see Figure 5.2. At each step, the software will check the log file folder of each controller via FTP connection and synchronize the corresponding folders on the data server.



```
config - Notepad
File Edit Format View Help
index,intersection id,ip addr,FTP Account Name, FTP Account PSW;
1,0501,134.84.148.10,econolite,ecpi2ecpi;
1,0502,134.84.148.11,econolite,ecpi2ecpi;
```

Figure 5.1 Configuration file

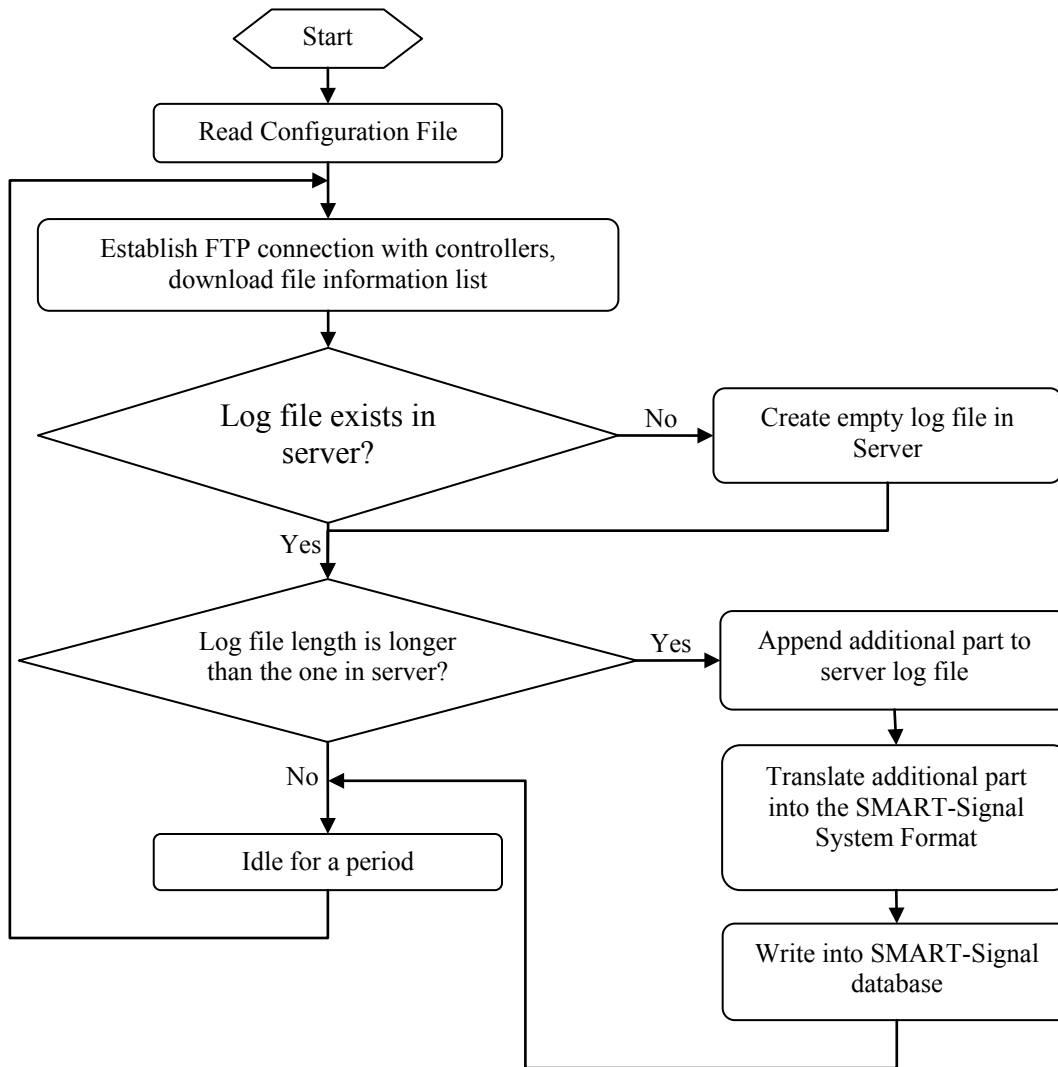


Figure 5.2 Flow chart for data retrieving software

The raw data within the Econolite ASC/3 controller are stored in hourly log files. Each log file starts with a header describing setting information of controller such as IP address and logging start time (see Figure 5.3). After the header, all the events were recorded in the log file in a binary format. Every 4 bytes of data forms a data trunk which represents an event. The data format is shown in Table 5.1, where the first byte indicates the event code, the second byte specifies the parameter and the last two bytes describes the time of the event. After translating the binary format into the decimal format, a data trunk example can be seen in Figure 5.4. Then, we need to map the event codes into specific event types happening at intersection and store the event information into the SMART-Signal database. The event code interpretation associated with signal changes and detector actuations is listed in Table 5.2..

Table 5.2 Econolite ASC/3 event code interpretation

Event Code	Event Descriptor	Parameter
Active Phase Events:		
0	Phase On	Phase # (1-16)
1	Phase Begin Green	Phase # (1-16)
7	Phase Green Termination	Phase # (1-16)
8	Phase Begin Yellow Clearance	
9	Phase End Yellow Clearance	
Detector Events:		
81	Detector Off	DET Channel # (1-64)
82	Detector On	DET Channel # (1-64)
89	PedDetector Off	DET Channel # (1-16)
90	PedDetector On	DET Channel # (1-16)

5.2 Implementation on TH 55

The new version of the SMART-Signal system was implemented in the 10 intersections along TH 55, Minneapolis, MN. Before the implementation, the list of IP addresses for each controller is needed for configuration of the data retrieving software. Currently, the implementation is based on the firmware version **02.51.00** of ASC/3 controllers. If the data server is directly connected to the field cabinet, i.e., without firewall in between, all the implementation can be completed on the server side. Follow the procedures below,

- 1) Enable the data logging function of ASC-3 controllers
- 2) On a new ASC/3 controller, the default value of **asc3DataLogEnable** is OFF and the default value of **asc3DataLogCircularBufferEnable** is ON. Upon power-up, the controller will not be logging any data. To enable the data logging function, run the **ASC3_SNMP_Util** toolbox on the data server and connect to specific controller by specifying the IP address. Through the toolbox, set the values of **asc3DataLogEnable** and **asc3DataLogCircularBufferEnable** to 1.
- 3) Install the SMART-Signal software packages on the data server and make necessary changes to configuration files.
- 4) Start the SMART-Signal data-retrieving software to retrieve data from field controllers.
- 5) Start other SMART-Signal programs to calculate performance measures.

The real-time performance of the 10 intersections along TH 55 can be monitored on the SMART-Signal website (<http://dotapp7.dot.state.mn.us/smartsignal/>) and Figure 5.5 gives an example of intersection level of service for the 10 intersections.

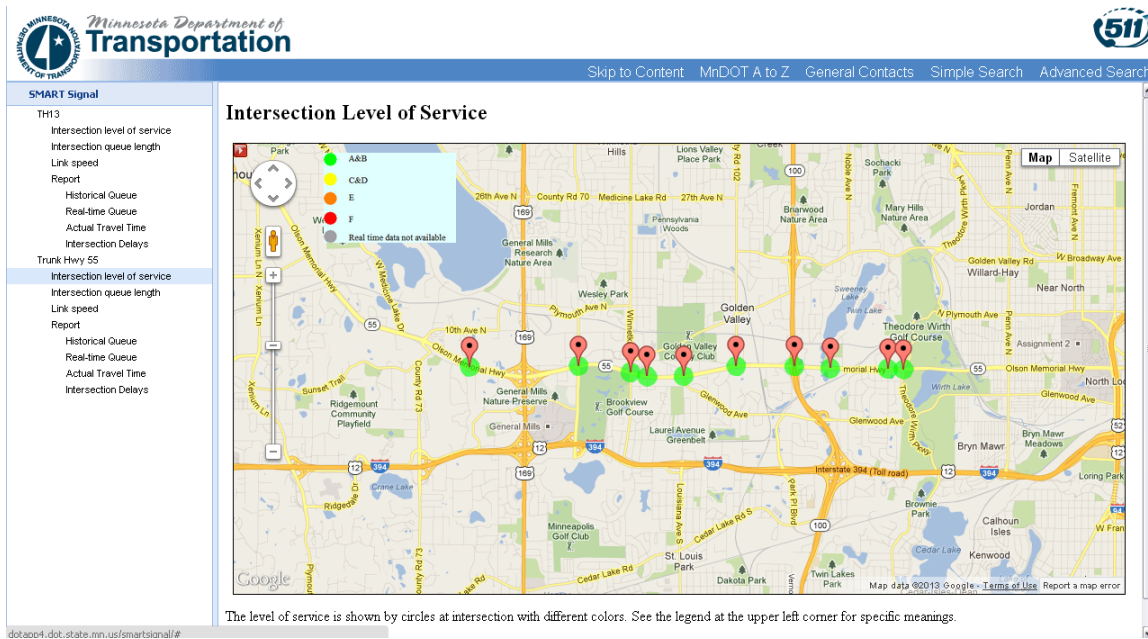


Figure 5.5 Intersection level of service along TH 55 (Map Source: Google Maps)

Chapter 6. Concluding Remarks

Due to the increasing traffic demand but limited facility capacities, traffic congestion has become an increasingly severe problem for metropolitan areas not only in the United States but also around the world. How to efficiently and effectively manage traffic during peak hours or non-recurrent congestion periods appears to be a challenging task for researchers and practitioners. The Integrated Corridor Management (ICM) approach has drawn additional attention in recent years because it is believed to be a promising tool to manage urban traffic congestion. In practice, the application of the ICM concept is still at the very early stage and most of the work remains at the policy research level.

In this project, a simple but effective maximum flow based control model was developed to handle oversaturated traffic conditions at signalized arterials. The model is built upon the *oversaturation severity index (OSI)*, which indicates not only the level of congestion, but also the reason for oversaturation. To solve the model, a Forward-Backward Procedure (FBP) is proposed and it is mathematically proven that it generates the optimal solution to the model. Based on the arterial control model, an integrated control model was proposed to manage network congestion. Through diversion control, the model tries to fully utilize the available capacity along parallel routes. The impact of the diverted traffic is specifically considered, especially for signalized arterials, so the caused congestion can be reduced or eliminated by proper adjustment of signal timings. This model does not rely on time-dependent traffic demand as model inputs and it is ready to be implemented at typical parallel traffic corridors where the standard detection system is available. With the extremely low computation burden, the model is very suitable for on-line applications. We have tested the performance of the proposed model using microscopic traffic simulation in the I-394 and TH 55 corridor in Minneapolis. The results indicate that the proposed model can significantly reduce network congestion and make traffic much smoother, which can be reflected by the huge improvement in network performance measures, such as average delay per vehicle, average number of stops per vehicle and average speed.

A new version of the SMART-Signal data collection software was also developed as part of this project. The software can directly retrieve event-based traffic data from the Econolite ASC/3 controllers without additional hardware instrumentation. The software was implemented in 10 intersections along TH 55 in Minneapolis.

The ICM control system developed in this project has a very promising future for real field implementation and we look forward to testing the field performance of the proposed approach in future projects.

References

- Abdelghany K. F.,Valdes, D.M., Abdelfatah, A.S. and Mahmassani, H.S. 1999. “Real-Time Dynamic Traffic Assignment and Path-Based Signal Coordination; Application to Network Traffic Management”. *Transportation Research Record* 1667, 67-76.
- Adler, J.L. and Blue, V.J. 2002. “A cooperative multi-agent transportation management and route guidance system”. *Transportation Research Part C* 10(5-6), 433-454.
- Ahn, S., Bertini, R.L., Auffray, B., Ross, J.H. and Eshel, O. 2007. “Evaluating benefits of system wide adaptive ramp-metering strategy in Portland, OR”. *Transportation Research Record* 2012, 47-56.
- Alexiadis, V. 2008. “Integrated Corridor Management Analysis, Modeling, and Simulation Results for the Test Corridor”. Technical report, Federal Highway Administration, Washington, D.C.
- Alm, E., Lingham, V., Benouar, H., Ban, X.J. and Chu, L. 2008. “An Integrated Methodology for Corridor Management Planning”. Transportation Research Board 87th Annual Meeting.
- Ben-akiva, M., Bierlaire, M., Bottom, J., Koutsopoulos, H. and Mishalani, R. 1997. “Development of a Route Guidance Generation System for Real-Time Application”. 8th IFAC Symposium on Transportation Systems, Chania, Creta, Greece
- Chang, G.L., Ho, P.K. and Wei, C.H. 1993. “A dynamic system-optimum control model for commuting traffic corridors”. *Transportation Research Part C* 1(1), 3-22.
- Chang, K. and Stephanedes, Y. J. 1993. “Optimal control of freeway corridors”. *ASCE Journal of Transportation Engineering* 119(4), 504-514.
- Cronin, B., Mortensen, S., Sheehan, R. and Thompson, D. 2010. “Integrated Corridor Management”. *Public Roads* 74(3), 6-11.
- Dell’Olmo, P. and Mirchandani,P.B. 1995. “REALBAND: An Approach for Real-Time Coordination of Traffic Flows on a Network”. *Transportation Research Record* 1494, 106-116.
- Diakaki, C., Papageorgiou, M., McLean, T. 2000. “Integrated traffic-responsive urban corridor control strategy in Glasgow, Scotland: Application and Evaluation”. *Transportation Research Record* 1727, 101-111.
- FHWA, 1988. TRANSYT User’s Manual (Release 6), US Department of Transportation, Federal Highway Administration, Washington, D.C.
- FHWA. 2005.Integrated Corridor Management System (ICMS) Work Plan. http://www.itsdocs.fhwa.dot.gov/icms/icms_workplan.htm. Accessed 2005.

- FHWA. 2007. Managing Congestion with Integrated Corridor Management. FHWA-JPO-07-025.
- FHWA, 2008a. Concept of Operations for the US-75 Integrated Corridor in Dallas, Texas. FHWA-JPO-08-004.
- FHWA, 2008b. Concept of Operations for the I-15 Corridor in San Diego, California. FHWA-JPO-08-009.
- Gartner, N. H. 1983. OPAC: A demand-responsive strategy for traffic signal control. U.S. Transportation Research Record 906.
- Haj-Salem, H., Papageorgiou, M. 1995. Ramp metering impact on urban corridor traffic: Field results. *Transportation Research Part A* 29(4), 303-319.
- Hamdar, S.H., Eisenman, S.M., Mahmassani, H.S. 2006. "Evaluation of Operational Strategies for Integrated Corridor Management". Transportation Research Board 85th Annual Meeting, Paper number 06-3096.
- Hasan M., Jha, M., Ben-Akiva.,M. 2003."Evaluation of ramp control algorithms Using Microscopic Simulation". *Transportation Research Part C* 10(3), 229-256.
- Hawas, Y. and Mahmassani, H.S. 1995. "A Decentralized Scheme for Real-time Route Guidance in Vehicular Traffic Networks". Intelligent Transport Systems World Congress, Yokohama, Japan.
- Head, K.L., Mirchandani, P.B. and Sheppard,D. 1992. "Hierarchical Framework for RealTime Traffic Control". *Transportation Research Record* 1360, 82-88.
- Hunt P.B., Robertson D.I. and Bretherton R.D., 1982. "The SCOOT on-line traffic signal optimization technique (Glasgow)". *Traffic Engineering & Control* 23(4) 190-192.
- Kattan, L., Nurul Habib, K. M., Nadeem, S. and Islam, T. 2010. "Modeling Travelers' Responses to Incident Information Provided by Variable Message Signs in Calgary, Canada". *Transportation Research Record*, vol.2185, pp 71-80.
- Kotsialos, A.,Papageorgiou, M.,Mangeas, M. and Haj-Salem, H. 2002."Coordinated and integrated control of motorway networks via non-linear optimal control". *Transportation Research Part C* 10(1), 65-84.
- Lee, C., Ran, B., Yang, F. and Loh, W. 2010. "A Hybrid Tree Approach to Modeling Alternate Route Choice Behavior with Online Information". *Journal of Intelligent Transportation Systems*, vol. 14, pp. 209-219.
- Liu, Y., Chang, G.L. 2010. "A generalized corridor diversion control model for freeway incident management". 89th TRB Annual Meeting Proceedings, CDROM, 10-2969.

- Liu, Y. and Chang, G. and Yu, J. 2011. "An Integrated Control Model for Freeway Corridor Under Nonrecurrent Congestion". *IEEE Transactions on Vehicular Technology* 60(4), 1404-1418.
- Ma, J., Nie, Y. and Zhang, H.M.,2007. "Solving the Integrated Corridor Control Problem Using Simultaneous Perturbation Stochastic Approximation". Transportation Research Board 86th Annual Meeting.
- MacCarley, C.A., Mattingly, S.P., McNally, M.G., Mezger, D. and Moore, J.E. 2002. "Field Operational Test of integrated freeway ramp metering/arterial adaptive signal control lessons learned in Irvine, California". *Transportation Research Record* 1811, 76-83.
- Mahmassani, H.S., 2001. "Dynamic network traffic assignment and simulation methodology for advanced system management applications". *Networks and Spatial Economics* 1(3/4), 267-292.
- Mahmassani, H.S., Sbayti, H.A. Zhou, X.,2004. "DYNASMART-P Version 0.930.7 User's Guide". Prepared for U.S. Department of Transportation.
- Messmer, A. and Papageorgiou, M. 1995. "Route Diversion Control in Motorway Networks via Nonlinear Optimization". *IEEE Transactions on Control Systems Technology*, Vol. 3, No. 1, 144-154.
- Minciardi, R. 2001. "A Decentralized Optimal Control Scheme for Route Guidance in Urban Road Networks". *IEEE Intelligent Transportation Systems Conference Proceedings*, Oakland, CA.
- Papageorgiou, M. 1990. "Dynamic modeling, assignment, and route guidance in traffic networks". *Transportation Research Part B* 24, 471-495.
- Papageorgiou, M. 1995. "An integrated control approach for traffic corridors". *Transportation Research Part C* 3(1),19-30.
- Papageorgiou, M.,Diakaki,C. and Mclean, T., 1997." Simulation studies of integrated corridor control in Glasgow". *Transportation Research Part C* 5(3), 211-224.
- Papageorgiou, M., Haj-Salem,H. and Middelham, F. 1998. "ALINEA Local Ramp Metering: Summary of Field Results". *Transportation Research Record* 1603, 13-20.
- Papageorgiou, M., Diakaki, C., Dinopoulou, V., Kotsialos, A. and Wang, Y. B. 2003. "Review of road traffic control strategies". *Proceedings of the IEEE*, vol. 91, pp. 2043-2067.
- Pavlis, Y. and Papageorgiou, M. 1999. "Simple Decentralized Feedback Strategies for Route Guidance in Traffic Networks". *Transportation Science*. Vol. 33, No. 3, 264-278
- Paz, A., Peeta, S. 2009. "Behavior-consistent real-time traffic routing under information provision". *Transportation Research Part C* 17(6), 642-661.

- Peeta, S. and Gedela, S., 2001. “Real-time variable message sign-based route guidance consistent with driver behavior”. *Transportation Research Record*, vol. 1752, pp. 117-125.
- Peeta, S., Ramos, J. L. and Pasupathy, R. 2000. “Content of variable message signs and on-line driver behavior”. *Transportation Research Record*, vol. 1725, pp. 102-108.
- Perugu, H.C., Wei, H.,Kashayi, N.C., 2007. “Heuristic System Model and Algorithms for Integrated Urban CorridorTraffic Control System with Access Management Measures”. Transportation Research Board 86th Annual Meeting, 07-1515.
- Reiss, R., Gordon, R., Neudorff, L. and Harding, J. 2006. “Integrated Corridor Management Phase I Concept Development and Foundational Research: Task 3.1 Develop Alternative Definitions”. Report No. FHWA-JOP-06-034. United States Department of Transportation ITS Joint ProgramOffice, Federal Highway Administration, Federal Transit Administration. April 11, 2006.
- Robertson, D.I. and Vincent, R.A. 1974. “Bus priority in a network of fixed-time signals”. Proceedings of 6th International Symposium on Transportation andTraffic Theory, 285-306. Sydney.
- Robertson, D. Bretherton,R. 1991. “Optimizing networks of traffic signals in real time — The SCOOT method”. *IEEE Transaction on Vehicular Technology* 40(1), 11-15.
- Sen, S. and Head, L. 1997. “Controlled optimization of phases at an intersection”. *Transportation Science*, vol. 31, pp. 5-17.
- Stephanedes, Y.J. and Kwon, E. 1993. “Adaptive demand-diversion prediction for integrated control of freeway corridors”. *Transportation Research Part C* 1(1), 23-42.
- Sussman, J.M., et al. 2000. “What Have We Learned About ITS?” Report No. FHWA-OP-01-006. Federal Highway Administration, U.S. Department of Transportation. Washington, D.C.
- Tian, Z. 2007. “Modeling and Implementation of an Integrated Ramp Metering-Diamond Interchange Control System”. *Journal of Transportation Systems Engineering and Information Technology* 7(1), 61-69.
- Van Aerde, M., Yagar, S. 1988a. “Dynamic integrated freeway/traffic signal networks: Problems and proposed solutions”. *Transportation Research Part A* 22(6), 435-443.
- Van Aerde, M., Yagar, S. 1988b. “Dynamic integrated freeway/traffic signal networks: A routing-based modeling approach”. *Transportation Research Part A* 22(6), 445-453.
- Van den Berg, M., De Schutter, B.,Hegyí, A. and Hellendoorn, J. 2001 “Model predictive control for mixed urban and freeway networks”. *Transportation Research Record* 1748, 55-65.

- Wang, Y. and Papageorgiou, M. 2002. "A Predictive Feedback Routing Control Strategy for Freeway Network Traffic". Proceedings of the American Control Conference, Anchorage, AK.
- Wattleworth, J. A. 1965 "Peak-period analysis and control of a freeway system," Highway Res. Board, Washington, D.C., *Highway Research Record* 157.
- Wu, J. and Chang, G.L.1999. "An integrated optimal control and algorithm for commuting corridors". *International Transactions in Operational Research* 6(1), 39-55.
- Yang, H. and Yagar, S. 1995. "Traffic assignment and signal control in saturated road networks". *Transportation Research Part A* 29(2), 125-139.
- Zhang, H.M. and Recker, W.W. 1999. "On optimal freeway ramp control policies for congested traffic corridors". *Transportation Research Part B* 33, 417-436.