USDOT Region V Regional University Transportation Center Final Report

NEXTRANS Project No. 019FY02

OPTIMAL SIGNAL TIMING DESIGN FOR URBAN STREET NETWORKS UNDER USER EQUILIBRIUM BASED TRAFFIC CONDITIONS

By

Dr. Zongzhi Li
Associate Professor
Illinois Institute of Technology
lizz@iit.edu

and

Dr. Lili Du
Assistant Professor
Illinois Institute of Technology
ldu3@iit.edu

and

Yi Liu, Ph.D.
Graduate Research Assistant
Illinois Institute of Technology
yliu136@hawk.iit.edu

DISCLAIMER

Funding for this research was provided by the NEXTRANS Center, Purdue University under Grant No. DTRT12-G-UTC05 of the U.S. Department of Transportation, Office of the Assistant Secretary for Research and Technology (OST-R), University Transportation Centers Program. The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

USDOT Region V Regional University Transportation Center Final Report

TECHNICAL SUMMARY

NEXTRANS Project No. 019FY02 Final Report, September 20, 2016

Title

OPTIMAL SIGNAL TIMING DESIGN FOR URBAN STREET NETWORKS UNDER USER EQUILIBRIUM BASED TRAFFIC CONDITIONS

Introduction

In the ever-growing travel demand, traffic congestion on freeways and expressways recurs more frequently at a higher number of locations and for longer durations with added severity. This becomes especially true in large metropolitan areas. Particular to the urban areas, excessive crowdedness caused by inefficient traffic control also results in urban street networks operating in near or over-saturated conditions, leading to unpleasant travel experience due to long delays at intersections. As a consequence, the recurrent traffic congestion on roadway segments and vehicle delays at intersections inevitably compromise energy efficiency, traffic mobility improvement, safety enhancement, and environmental impacts mitigation. All too often, neither restraining travel demand nor expanding system capacity is desirable and practical. Conversely, effectively utilizing the capacity of the existing transportation system has been increasingly thought of as the solution to congestion relief. With respect to the urban street networks, developing effective means for urban intersection signal optimization becomes essential to reduce intersection delays.

Traffic signal control is used to determine who has the right of way at a signalized intersection and also able to control the flow patterns of traffic through the intersection. Early contributions in this area were mainly focusing on optimize signal settings, such as the total cycle time and the green splits, for a single isolated intersection (Webster, 1958). Such approaches could surely reduce the vehicle delays at single intersection, and be stretched to the entire network by applying the techniques to every intersection in the network. However, coordination between traffic signals in close proximity and their mutual effect on the network traffic assignment are not considered.

Conventional fixed time signal plan optimization strategies, as mentioned earlier, use historical traffic data and assume that traffic flows will remains unchanged after the implementation of new signal plans. Traffic flows were assumed to be given and invariable, but, in fact, when signal timings change, travel times for certain or all travel routes will be different, which definitely makes drivers in the network to adjust their choice of travel paths to destinations, and result in changes of traffic flows in the network. Then, new optimal signal settings are always required if they were treated independently with traffic flows. In an attempt to maintain the interdependency between traffic assignment and signal control, it was put forward that impacts of signal settings on the traffic flows should be considered by combining traffic control and route choice (Allsop 1974).

There are two different ways to solve this problem: the iterative optimization and assignment procedure and the combined optimization and assignment model. This study falls in the latter approach.

The iterative optimization and assignment procedure is to alternatively update signal timings with fixed traffic flows and solve the traffic equilibrium problem for the new signal settings until a mutually consistent (MC) solution is gained. This approach has the advantage that it actually solves the traffic assignment problem and signal timing optimization problem separately, using traditional traffic assignment and signal timing optimization techniques. Also, it can be applied on large networks much more easily compared to combined optimization and assignment model approach. But it has been pointed out theoretically and empirically by Dickson (1981) that this approach cannot guarantee to converge even to a local optimum, and also may lead to an increase in total delay over the network rather than a decrease. And this the main reason that leads this study to the combined signal timing optimization and traffic assignment model approach.

The combined signal timing optimization and traffic assignment model seeks optimal signal settings such that one or more system performance measures like the total travel time or average delay are minimized, while the driver's routing is simultaneously ascribed by a traffic equilibrium model. This combined problem is an instance of the network design problem (NDP), which is concerned with improving an existing network, meanwhile considering the user's response to the change. A bi-level structure can be employed to model this combined problem in which signal timing optimization is regarded as the upper level problem while user equilibrium traffic assignment is regarded as the lower level problem. Two major difficulties that involved in this approach need to be mentioned. One of them, which is with respect to problem solving, is that the problem is generally hard to solve because of the high complexity which comes from the non-convexity of objective functions and constraints at both level. Most previous studies on this approach focused on seeking an efficient algorithm, which was capable of finding a local optima or near-global optima of the signal setting variables in the upper level and simultaneously finding the user-decided flow pattern for the lower level. Another major difficulty of the combined problem approach, which is much less extensively studied, is the integration of time delay at signalized intersection to the bi-level optimization model. Some of the existing works, which are analytical-based, use an oversimplified assumptions of the traffic signal control system, and apply it to a small sample network. While other methods, most recent works, which were simulation-based, require an existing simulation model of the network such as TRANSYT to evaluate the performance of the system with different signal settings, and demands extensive data and work to build such simulation model before optimization. It is a costly and time consuming work when dealing with real world problems, comparing with other methods that uses analytical mathematical expression to formulate the system, and is generally not possible for states and agencies that do not maintain rich data on travel demand, facility preservation, traffic operations, data processing and preparation capacity, and have high performance computing facilities. The above-mentioned shortcomings all constraint potential employment of the methods for a real world network.

The above mentioned shortcomings motivate the author to address the combined optimization and assignment problem analytically as a rigorous Mathematical Problem with Equilibrium Constraints (MPEC) general assumptions and formulations, meanwhile, model and evaluate traffic signal controlled systems accurately, and also makes it well applicable to real world problems. Delay calculation and static user equilibrium will all be formulated as Variational Inequality constraints which allow the state-of-the-art MPEC solver, GAMS/NLPEC, to be employed for solving the problem for a local optimal effectively and efficiently.

Moreover, as one variable of the signal settings that can be adjusted to potentially improve the effectiveness of traffic signal plans, signal phasing design has received little attention from researchers. In bandwidth maximization approach, it was indicated that left-turn sequence had significant effectiveness. Tian et al. (2008) indicated that lead-lag phasing showed clear advantages over other phasing sequence in maximizing progression bandwidth. Meanwhile, various signal phasing designs provide different behaviors in terms of delays for different approach. For instance, comparing to lag left-turn phasing, lead left-turn phasing, which has a protected left-turn phase prior to through phase, can reduce the average signal control delays for corresponding left-turn traffic by reducing the continuous queuing time, which, as a consequence, reduces the maximum queue length. While, this also results in more continuous queuing time assigned to the opposing through traffic. Main reason that limits adding signal phasing to the delay-based signal optimization programs is the computational infeasibility that appeared (Cohen and Mekemson, 1985). Based on the MPEC model proposed earlier, the author attempts to develop an enhanced model that takes different signal phasing design into account. To model the selection of signal phasing designs, integer variables will be required in the enhanced model. However, GAMS/NLPEC solver has its limitation when dealing with integer programing, and it was built to solve problems with continuous variable only. Therefore, different solution algorithms will be proposed and attempt to solve this enhanced model.

In general, the author introduces a new methodology in this study that addresses the combined signal timing optimization and traffic assignment problem analytically with general assumptions and formulations, which models and evaluates traffic signal controlled systems accurately, and is also well applicable to real world problems. In the proposed method, urban network traffic signal timing optimization and traffic flow equilibrium problem are considered as a bi-level optimization problem. A basic model is proposed firstly, which attempt to minimize system total travel time by optimizing signal green splits. HCM 2010 delay method, which is one of the most up-to-date time-dependent stochastic delay models, is employed as intersection delay estimation method in the model, and is formulated as Variational Inequality constraints, what allow the state-of-the-art MPEC solver, GAMS/NLPEC, to be employed for solving the problem for a local optimal effectively and efficiently. A small sample network and a real world network in the densely populated City of Chicago area are used to test the capability of the model and the applicability to real world case in urban area.

Furthermore, in order to import more reality to the basic model and also consider the potential system benefit that comes from different signal phasing design, an enhanced model is developed based on the basic model by employing integer and binary variables. Then, the enhanced model belongs to a new class of challenging optimization problems, namely Mixed-Integer Nonlinear Programming (MINLP) with Complementarity Constraints. Formulating the problem with binary variables allows for the selection of proper phasing design. Heuristic solution algorithms are proposed and tested in a small sample network.

Findings

In this study, intersection control delay calculation method introduced in HCM 2010 has been employed in a combined optimization problem for area traffic signal control and network traffic assignment, and formulated as Variational Inequality (VI) constraints in the basic MPEC model. It allows the proposed method to accurately model and estimate the intersection control delay of various type of movements in real world scenarios such as those with multiple green phases and multiple control methods (protected, permitted, or mixed) without the use of simulation-based traffic model. The combined problem was formulated as mathematical programming with equilibrium constraints (MPEC) and solved by using GAMS/NLPEC solver which reformulates and solves the MPEC problem as standard nonlinear programming (NLP).

The basic MPEC model was applied on an experimental 4-intersection network and a real world problem with 13 signalized intersection in the City of Chicago urban area. Different phasing plans were adopted in the experimental network, and three traffic loads were tested as different cases from low traffic demand condition case (with intersection V/C around 30%) to high traffic demand condition case (with intersection V/C around 150%). Comparing the optimization results of the proposed model with the optimization results by using Synchro with the same initial traffic assignment, improvements in both total intersection control delay and total travel cost were observed in all three cases, and they varied significantly. Small improvement, 2.55% in total travel time reduction, was obtained in the low demand case, and large improvement, 14.54% in total travel time reduction, was showed in the medium demand case which has near capacity traffic loads at signalized intersections. After the optimization, drivers tended to switch their route from intersections with protected only left-turn phasing to intersections with protected-permitted left-turn phasing and split phasing, where more left turn traffic would better utilize the intersection capacity. Comparing with the protected left-turn only phasing, protected-permitted left-turn phasing and split phasing had relatively more capacity without occupying the green time for other phases.

For the real world problem, named as network two, two different OD demands generated by Chicago TRANSIMS microscopic traffic simulation model were tested. In the case with AM peak traffic demand, which is roughly 30% V/C, 11.23% total travel time reduction was obtained from the proposed method when compared with Synchro optimization result, and almost all of the travel time reduction was contributed by reduction in intersection control, 20.53%, in that network total link travel time remained basically the same with 0.28% increase. Under similar demand condition, 30% V/C ratio, the basic MPEC model tend to be more applicable and beneficial in larger network than small network with limited paths and intersections. Besides, it was also observed that changes in the traffic routing was the main reason and power that caused the improvement in system performance, and is also the major difference between Synchro and the basic MPEC model proposed in this study. However, in the case with off peak traffic demand, although the significance of result comparison was lost because of the bad optimization results from Synchro, it was still able to present the capability of the proposed model when dealing with extremely low demand situation.

Furthermore, in order to import more reality to the basic model and consider the potential system benefit that comes from different signal phasing designs, an enhanced model is developed based on the basic MPEC model by employing binary variables to make selection of optimal signal phasing plans from pre-defined candidates. The enhanced model belongs to a new class of challenging optimization problems, namely Mixed-Integer Nonlinear Programming (MINLP) with Complementarity Constraints. Formulating the problem with binary variables allows for the selection of proper phasing design, however, also increase the difficulty to solve the problem. As preliminary solution attempts, two heuristic solution algorithms, GA method and EA method, are proposed.

Both GA and EA solution method were implemented on the test network to verify the feasibility of the solution methods. In the network, one lead-lead left turn phasing and one split phasing designs were prepared as candidates for intersection 2 and 3, respectively. In total, 4 different combination of phasing plans were available in the problem.

Among two preliminary solution methods, GA failed to provide valid nor optimal solution within valid running period. While EA method, which highly relies on the basic MPEC model, provided optimal results when keeping original phasing at intersection 2 unchanged and replacing the split phasing at intersection 3 with normal lead-lead phasing. Comparing with the optimization results of the original phasing plans, 3.58%, 17.46%, and 11.77% reduction in network total cost were observed under low, medium, and high traffic demand conditions, respectively. Similar to previous cases, all reduction came from the improvement at signalized intersections, particularly, from intersection 3. The results strongly supported our assumption that adding phasing design as a variable in the model would further generate potential improvement in the system.

Recommendations

The application of the basic MPEC model, along with the solution method, does not require extensive data collection, preparation, and computational efforts as compared to the methods that rely on simulation-based traffic models to evaluate the performance of traffic signals. This gives it potentially greater applications to agencies that do not maintain rich data on travel demand, facility preservation, traffic operations, data processing and preparation capacity, and high performance computing facilities. However, current solution method relies on a good initial point to obtain an acceptable optimization result, and it would be useful to develop a better method to find a good initial point or initial feasible solution as future work.

For the enhanced model, an efficient solution algorithm is still under development. Both of the proposed preliminary solution methods have their limitations and required more research. Looking for an alternative of SUE, which allows more tolerance when locating feasible solutions, could be a future research direction for the GA method approach. Meanwhile, for EA method approach, a reduction method, which is able to effectively reduce the size of candidate phasing design combinations without losing solution optimality, are also needed to improve method's efficiency.

Contacts

For more information:

Dr. Zongzhi Li, Ph.D., Associate Professor
Illinois Institute of Technology
3201 South Dearborn Street, AM Hall, Room 211
Chicago, IL 60616
Phone: (312) 567-3556
Fax: (312) 567-3519
Email: lizz@iit.edu

NEXTRANS Center
Purdue University - Discovery Park
3000 Kent Ave.
West Lafayette, IN 47906
Phone: (765) 496-9724
Email: nextrans@purdue.edu
www.purdue.edu/dp/nextrans

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviations	Term
DTA	Dynamic Traffic Assignment
EA	Enumerative Algorithm
GA	Genetic Algorithm
GAMS	General Algebraic Modeling System
HCM	Highway Capacity Manual
LT	Left Turn
MC	Mutually Consistent
MINLP	Mixed-Integer Nonlinear Programming
MPEC	Mathematical Programming with Equilibrium Constraints
NDP	Network Design Problem
NLP	Non-Linear Programming
OD	Origin/Destination
SCOOT	Split Cycle Offset Optimisation Technique
SUE	Static User Equilibrium
TRANSIMS	TRansportation ANalysis and SIMulation System
TRANSYT	TRAffic Network StudY Tool
TRT	Through and Right Turn
V/C	Volume/Capacity

LIST OF SYMBOLS

Symbol	Definition
$\boldsymbol{N}$	the set of nodes
$\boldsymbol{A}$	the set of road links (links)
$\boldsymbol{N}_s$	a subset of $\boldsymbol{N}$ that includes all the fixed-time signal controlled intersections
$\boldsymbol{N}_k$	a subset of $\boldsymbol{N}_s$ that includes all signal controlled intersections that has cadidate phasing plans
$\boldsymbol{K}_n$	set of possible phasing designs for signalized intersection $n \in \boldsymbol{N}_k$, and its binary element k_n^m equals to 1 when phasing design m is selected for intersection n, 0 otherwise
$\boldsymbol{A}_n \subset \boldsymbol{A}$	the set of links that have a common head $n \in \boldsymbol{N}$
$\boldsymbol{I}$	the set of lane groups on link $a \in \boldsymbol{A}_n$ at any intersection $n \in \boldsymbol{N}'$
$\boldsymbol{L}_n$	the pre-defined signal phases at intersection $n \in \boldsymbol{N}'$
$\boldsymbol{O}$	the set of origin-destination (OD) pairs
$\boldsymbol{P}$	the set of paths
$\boldsymbol{P}_{nai} \subset \boldsymbol{P}$	the set of paths that traverse lane group i on link $a \in \boldsymbol{A}_n$ at intersection $n \in \boldsymbol{N}'$
α_o	the traffic demand between any OD pair o
$\boldsymbol{\Delta}$	link-path incidence matrix with elements $\delta_a^p = 1$, if path p traverse

link a, 0 otherwise

Θ OD-path incidence matrix with elements $\theta_o^p = 1$, if path p connect OD pair o, 0 otherwise

Γ lane group-path incidence matrix with elements $\gamma_{nai}^p = 1$, if path p traverse lane group i on link $a \in \boldsymbol{A}_n$ at intersection $n \in \boldsymbol{N}'$, 0 otherwise

T the analysis period

ω_n the cycle time at any signalized intersection $n \in \boldsymbol{N}'$

t_{lost} the total lost time per phase

$lnum_{nai}$ the number of lanes of lane group i on link $a \in \boldsymbol{A}_n$ at intersection $n \in \boldsymbol{N}'$

$s_{nai}^{l_n}$ the saturation capacity per lane in phase l_n

Z an incidence matrix with elements $\zeta_{nai}^{l_n',l_n} = 1$, if, for lane group i on link $a \in \boldsymbol{A}_n$ at intersection $n \in \boldsymbol{N}'$, phase l_n' is the previous phase of l_n $(l_n', l_n \in \boldsymbol{L}_n)$ in one cycle, 0 otherwise. Since the calculation of uniform delay is started from the first red light phase, and it is assumed that there is no initial queue, which means the queue length at the beginning of the first red light phase is 0. Thus, in order to achieve this, for the first red light phase of any lane group nai, all the elements are 0

$gmin_n^{l_n}$ minimum allowed green time associated with phase l_n at intersection $n \in \boldsymbol{N}'$

$gmax_n^{l_n}$ maximum allowed green time associated with phase l_n at intersection $n \in \boldsymbol{N}'$

$w_{nai}^{l_n}$ the queue change rate in phase l_n

τ_o the equilibrium travel cost between any OD pair o

f_p the traffic flow on any path p

h_a the travel time on any link a

c_p the total travel cost per vehicle for path p

$x_n^{l_n}$ the effective green for phase $l_n \in \boldsymbol{L_n}$ at signalized intersection $n \in \boldsymbol{N'}$

$Q_{nai}^{l_n}$ the queue length in number of vehicles at the end of each phase for lane group i on link $a \in \boldsymbol{A}_n$ at intersection $n \in \boldsymbol{N'}$ with. For the queue at the beginning of red phase, $Q_{nai}^{l_n} = 0$

f_{nai} the traffic flow rate on lane group i on link $a \in \boldsymbol{A}_n$ at intersection $n \in \boldsymbol{N'}$

$\bar{s}_{nai}$ the capacity of lane group i, which is the average capacity through whole cycle

$y_{nai}^{l_n}$ the queue change duration in phase l_n

d_{nai} the average control delay per vehicle

d_{nai}^1 the uniform delay

d_{nai}^2 the incremental delay

ABSTRACT

In the ever-growing travel demand, traffic congestion on freeways and expressways recurs more frequently at a higher number of locations and for longer durations with added severity. This becomes especially true in large metropolitan areas. Particular to the urban areas, excessive crowdedness caused by inefficient traffic control also results in urban street networks operating in near or over-saturated conditions, leading to unpleasant travel experience due to long delays at intersections. As a consequence, the recurrent traffic congestion on roadway segments and vehicle delays at intersections inevitably compromise energy efficiency, traffic mobility improvement, safety enhancement, and environmental impacts mitigation. All too often, neither restraining travel demand nor expanding system capacity is desirable and practical. Conversely, effectively utilizing the capacity of the existing transportation system has been increasingly thought of as the solution to congestion relief. With respect to the urban street networks, developing effective means for urban intersection signal optimization becomes essential to reduce intersection delays.

Conventional signal timing optimization methods use historical traffic data and assume that traffic flows will remains unchanged after the implementation of new signal timing plans. Traffic flows are assumed to be constant, but in fact, when signal timing plans change, travel times for some travel routes will alter, which requires drivers in the network to adjust their choice of travel routes to arrive at the destinations, and result in

redistribution of traffic in the network. Therefore, the effects of interactions between signal timing plans and traffic flows need to be explicitly taken into consideration. This study introduces a new methodology that jointly considers signal timing optimization and traffic assignment in an overall analytical framework that contains model formulations under assumptions consistent with real world situations. Such a framework is well suited for applications in real world cases. Specifically, the overall optimization framework is formulated as a bi-level optimization problem. In the proposed basic model, at the upper level, a traffic signal timing optimization problem for urban network is introduced to minimize system total travel time by optimizing signal green splits. At the lower level, a static user equilibrium problem is formulated for networkwide traffic assignment. In the vehicle delay estimation, the time-dependent stochastic delay model in the 2010 Highway Capacity Manual (HCM 2010) is employed and formulated as Variational Inequality constraints, what allow the state-of-the-art MPEC solver, GAMS/NLPEC, to solve the problem for a local optimal effectively and efficiently. The bi-level optimization model is first tested using a small network (the test network) and a computational experiment using a subarea network in the Chicago central district is conducted to assess the practicality of the model formulation in real world applications.

In order to import more reality to the basic model and also consider the potential system benefit that comes from different signal phasing design, an enhanced model is developed based on the basic model by employing integer and binary variables. Formulating the problem with binary variables allows for the selection of proper phasing design. Heuristic solution methods are proposed and tested using the test network.

CHAPTER 1. INTRODUCTION

In the ever-growing travel demand, traffic congestion on freeways and expressways recurs more frequently at a higher number of locations and for longer durations with added severity. This becomes especially true in large metropolitan areas. Particular to the urban areas, excessive crowdedness caused by inefficient traffic control also results in urban street networks operating in near or over-saturated conditions, leading to unpleasant travel experience due to long delays at intersections. As a consequence, the recurrent traffic congestion on roadway segments and vehicle delays at intersections inevitably compromise energy efficiency, traffic mobility improvement, safety enhancement, and environmental impacts mitigation. All too often, neither restraining travel demand nor expanding system capacity is desirable and practical. Conversely, effectively utilizing the capacity of the existing transportation system has been increasingly thought of as the solution to congestion relief. With respect to the urban street networks, developing effective means for urban intersection signal optimization becomes essential to reduce intersection delays.

Traffic signal control is used to determine who has the right of way at a signalized intersection and also able to control the flow patterns of traffic through the intersection. Early contributions in this area were mainly focusing on optimize signal settings, such as the total cycle time and the green splits, for a single isolated intersection (Webster, 1958). Such approaches could surely reduce the vehicle delays at single intersection, and be stretched to the entire network by applying the techniques to every intersection in the network. However, coordination between traffic signals in close proximity and their mutual effect on the network traffic assignment are not considered.

Coordination between signalized intersections can provide multiple advantages along arterial streets (Homburger 1982). This is inevitable for early signal timing optimization methods in that vehicle arrivals at an intersection were assumed to be uniformly or randomly distributed. While in reality, traffic signals tend to group vehicles into a "platoon" so that for downstream intersections, most vehicles arrive as platoons with certain time interval that is highly related to the traffic signal at upstream intersections. Then, with a good coordination between signalized intersections, especially on arterials, continuous movement of vehicle platoons could be maintained. Since no existing mathematical model that minimizes overall travel times or delays can successfully model the platooning effect (Roess, Passas, & Mcshane, 2004), a surrogate approach that maximizes the bandwidth of the traffic progression is employed to solve the signal coordination problem. In this case, the signal settings are designed to maximize the width of continuous green bands along both directions of an arterial at certain speed. As a surrogate performance measure, maximizing the bandwidth can, of course, provide significant benefits to the traffic on target arterials, but these changes may lead to a worse traffic conditions on perpendicular roads, leading to a much smaller overall network benefit in terms of total travel time or delay reductions. Considering the signal coordination, strategies that optimized a group of signalized intersections were developed. Some of them, such as TRANSYT (Robertson 1969), from planning perspective, provided fixed time strategies using historical traffic flow data, while the others, such as SCOOT (Hunt, Robertson, & Bretherton, 1981), focused on real time operation that provides demand responsive strategies from real time traffic flow data. In this research, fixed time signal plans are studied.

Conventional fixed time signal plan optimization strategies, as mentioned earlier, use historical traffic data and assume that traffic flows will remains unchanged after the implementation of new signal plans. Traffic flows were assumed to be given and invariable, but, in fact, when signal timings change, travel times for certain or all travel routes will be different, which definitely makes drivers in the network to adjust their choice of travel paths to destinations, and result in changes of traffic flows in the network. Then, new optimal signal settings are always required if they were treated independently with traffic flows. In an attempt to maintain the interdependency between traffic assignment and signal control, it was put forward that impacts of signal settings on the traffic flows should be considered by combining traffic control and route choice (Allsop 1974).

There are two different ways to solve this problem: the iterative optimization and assignment procedure and the combined optimization and assignment model. This study falls in the latter approach.

The iterative optimization and assignment procedure is to alternatively update signal timings with fixed traffic flows and solve the traffic equilibrium problem for the new signal settings until a mutually consistent (MC) solution is gained. This approach has the advantage that it actually solves the traffic assignment problem and signal timing optimization problem separately, using traditional traffic assignment and signal timing optimization techniques. Also, it can be applied on large networks much more easily compared to combined optimization and assignment model approach. But it has been pointed out theoretically and empirically by Dickson (1981) that this approach cannot guarantee to converge even to a local optimum, and also may lead to an increase in total delay over the network rather than a decrease. And this the main reason that leads this study to the combined signal timing optimization and traffic assignment model approach.

The combined signal timing optimization and traffic assignment model seeks optimal signal settings such that one or more system performance measures like the total travel time or average delay are minimized, while the driver's routing is simultaneously ascribed by a traffic equilibrium model. This combined problem is an instance of the network design problem (NDP), which is concerned with improving an existing network, meanwhile considering the user's response to the change. A bi-level structure can be employed to model this combined problem in which signal timing optimization is regarded as the upper level problem while user equilibrium traffic assignment is regarded as the lower level problem. Two major difficulties that involved in this approach need to be mentioned. One of them, which is with respect to problem solving, is that the problem is generally hard to solve because of the high complexity which comes from the non-convexity of objective functions and constraints at both level. Most previous studies on this approach focused on seeking an efficient algorithm, which was capable of finding a local optima or near-global optima of the signal setting variables in the upper level and simultaneously finding the user-decided flow pattern for the lower level. Another major difficulty of the combined problem approach, which is much less extensively studied, is the integration of time delay at signalized intersection to the bi-level optimization model. Some of the existing works, which are analytical-based, use an oversimplified assumptions of the traffic signal control system, and apply it to a small sample network. While other methods, most recent works, which were simulation-based, require an existing simulation model of the network such as TRANSYT to evaluate the performance of the system with different signal settings, and demands extensive data and work to build such simulation model before optimization. It is a costly and time consuming work when dealing with real world problems, comparing with other methods that uses analytical mathematical expression to formulate the system, and is generally not possible for states and agencies that do not maintain rich data on travel demand, facility preservation, traffic operations, data processing and preparation capacity, and have high performance computing facilities. The above-mentioned shortcomings all constraint potential employment of the methods for a real world network.

The above mentioned shortcomings motivate the author to address the combined optimization and assignment problem analytically as a rigorous Mathematical Problem with Equilibrium Constraints (MPEC) general assumptions and formulations, meanwhile, model and evaluate traffic signal controlled systems accurately, and also makes it well applicable to real world problems. Delay calculation and static user equilibrium will all be formulated as Variational Inequality constraints which allow the state-of-the-art MPEC solver, GAMS/NLPEC, to be employed for solving the problem for a local optimal effectively and efficiently.

Moreover, as one variable of the signal settings that can be adjusted to potentially improve the effectiveness of traffic signal plans, signal phasing design has received little attention from researchers. In bandwidth maximization approach, it was indicated that left-turn sequence had significant effectiveness. Tian et al. (2008) indicated that lead-lag phasing showed clear advantages over other phasing sequence in maximizing progression bandwidth. Meanwhile, various signal phasing designs provide different behaviors in terms of delays for different approach. For instance, comparing to lag left-turn phasing, lead left-turn phasing, which has a protected left-turn phase prior to through phase, can reduce the average signal control delays for corresponding left-turn traffic by reducing the continuous queuing time, which, as a consequence, reduces the maximum queue length. While, this also results in more continuous queuing time assigned to the opposing through traffic. Main reason that limits adding signal phasing to the delay-based signal optimization programs is the computational infeasibility that appeared (Cohen and Mekemson, 1985). Based on the MPEC model proposed earlier, the author attempts to develop an enhanced model that takes different signal phasing design into account. To model the selection of signal phasing designs, integer variables will be required in the enhanced model. However, GAMS/NLPEC solver has its limitation when dealing with integer programing, and it was built to solve problems with continuous variable only. Therefore, different solution algorithms will be proposed and attempt to solve this enhanced model.

In general, the author introduces a new methodology in this study that addresses the combined signal timing optimization and traffic assignment problem analytically with general assumptions and formulations, which models and evaluates traffic signal controlled systems accurately, and is also well applicable to real world problems. In the proposed method, urban network traffic signal timing optimization and traffic flow equilibrium problem are considered as a bi-level optimization problem. A basic model is proposed firstly, which attempt to minimize system total travel time by optimizing signal green splits. HCM 2010 delay method, which is one of the most up-to-date time-dependent stochastic delay models, is employed as intersection delay estimation method in the model, and is formulated as Variational Inequality constraints, what allow the state-of-the-art MPEC solver, GAMS/NLPEC, to be employed for solving the problem for a local optimal effectively and efficiently. A small sample network and a real world network in the densely populated City of Chicago area are used to test the capability of the model and the applicability to real world case in urban area.

Furthermore, in order to import more reality to the basic model and also consider the potential system benefit that comes from different signal phasing design, an enhanced model is developed based on the basic model by employing integer and binary variables. Then, the enhanced model belongs to a new class of challenging optimization problems, namely Mixed-Integer Nonlinear Programming (MINLP) with Complementarity Constraints. Formulating the problem with binary variables allows for the selection of proper phasing design. Heuristic solution algorithms are proposed and tested in a small sample network.

This research is organized as follows: Chapter 1 provides an introduction of the background and motivations of this study, and is followed by a more detailed literature review Chapter 2. Chapter 3 introduces methodology and formulation of the proposed basic MPEC model, while in Chapter 4, computational experiments of the basic MPEC model on a small sample network and a real world problem will be presented. Similarly, in Chapter 5, the enhanced MINLP with complementarity constraints is developed and potential solution algorithms will be discussed. Computational experiments of the enhanced model on a sample small network is presented and discussed in Chapter 6. Finally, Chapter 7 summarizes the study and indicates the potential future research topic.

CHAPTER 2. LITERETURE REVIEW

As the first step of the research, literature review was conducted on existing methodologies for signal timing optimization and traffic assignment and intersection vehicle delay modelling as summarized in the following sections.

2.1 Signal Optimization and Traffic Assignment

Existing methodologies for signal timing optimization and traffic assignment can be categorized into two different approaches: the iterative optimization and assignment procedure and the combined optimization and assignment model.

For the iterative optimization and assignment procedure, Allsop and Charlesworth (1977) presented a mutually consistent calculation for the area traffic signal timing optimization problem and equilibrium traffic flows, in which signal timing optimization and traffic assignment procedure were treated alternatively and updated by solving signal timings with fixed traffic flows and solving the traffic equilibrium problem for the new signal settings.

Later on in the 1980s, Smith (1980, 1981a, 1981b) studied the existence and properties of the equilibrium between traffic control and traffic assignment from the perspective of a local control policy. Sheffi and Powell (1983) presented a mathematical programming formulation and a solution algorithm for a small network. In addition, they also provided two heuristic algorithms for a large-scale network application. More iterative optimization and assignment procedure based methods were proposed under static traffic assignment (Gartner, Gershwin, Little, & Ross, 1980; Cantarella, Improto, & Sforza, 1991; Gartner & Al-Malik, 1996).

Meanwhile, dynamic traffic assignment (DTA) was also been considered in some other works more recently. Abdelfatah and Mahmassani (1998) combined the signal control problem with the DTA problem by presenting a mathematical formulation and a simulation-based solution algorithm. With the help of transportation simulation tool, they applied their method on a realistic and moderately large network. In 2001, they extended this work by replacing the well-known Webster's formula by a simulation based signal timing optimization method to optimize signal control settings (Abdelfatah & Mahmassani 2001).

Another approach to address this problem is the bi-level programming methods for the combined optimization and assignment model. In the bi-level structure, the dependence of equilibrium flows on the decision variables is treated as a constraint of the signal optimization problem.

Heydecker and Khoo (1990) firstly presented a formulation of this combined problem as a bi-level problem and reported that, when compared with the iterative optimization and assignment procedure, the bi-level formulation improved the system performance in their sample network.

Yang and Yagar (1995) modeled the combined problem in saturated road networks as bi-level problem, considering the effects of travel routing from queuing, and determined equilibrium link flow and delay using the sensitivity analysis (SA) originally proposed by Tobin and Friesz (1988) and further developed by Friesz et al. (1990).

Meneguzzer (1995) employed diagonalization algorithm to solve the combined problem and successfully applied his work on a real suburban network in Chicago region.

In his work, the 1985 Highway Capacity Manual (HCM) methods were used in updating capacity and calculating link travel time and traffic control delays, and helped his work to be one of very few works that was applied on a real world case.

Chiou (1999) used projection method for local search and a heuristic approach for global search to solve the bi-level problem, in which the performance of the system, as a weighted sum of signal control delay and number of stops, was evaluated by use of the simulation-based traffic model, TRANSYT. Moreover, several enhanced heuristic solution algorithms were adopted.

Yin (2000) proposed a genetic algorithm (GA) based approach for bi-level programming models in transportation research. It was reported that the GA approach requires more computational efforts but avoids the complex computation of the derivatives of link performance functions for equilibrium network flows in SA approach. Ceylan and Bell (2004, 2005) utilized GA for combining the assignment software Path Flow Estimator (PFE) and TRANSYT. Moreover, Ceylan (2006) combined GA with TRANSYT Hill-Climbing optimization routine, and proposed a method for decreasing the search space to find optimal or near-optimal signal settings.

More recently, Ceylan and Ceylan (2012) presented a new hybrid Harmony Search and TRANSYT Hill-Climbing algorithm for signalized stochastic equilibrium transportation network.

Most of the recent works reviewed were focusing on extending the problem to dynamic traffic assignment approach and seeking better solution algorithms. When it comes to modeling and calculating the system performance measures to evaluate the system, most of the works tend to use a simulation-based traffic model such as TRANSYT, which required to be built before optimization. It will be a costly and time consuming work when dealing with real world problems, comparing with other methods that uses analytical mathematical expression to formulate the system, and is generally not possible for states and agencies that do not maintain rich data on travel demand, facility preservation, traffic operations, data processing and preparation capacity, and have high performance computing facilities. This also limits the potential of those methods to be applied on a real world problem. While using mathematical formulae to model a transportation system, it is difficult to find the best tradeoff between reality, optimality, and efficiency. In order to keep the reality of traffic operations and the reliability of model results, these models need to have high complexity with high degree of nonlinearity involved. As a consequence, they are generally hard to solve without a large number of assumptions and approximations, and yet with limited application to a real-world case in urban areas.

2.2 Intersection Delay Estimation

Among all the essential assumptions and approximations, the delay formula at signalized intersections is potentially the most important one. The same signal settings considered under different cost assumptions may provide totally different theoretical properties and result in completely different results. From the classic deterministic queuing model to the Shock wave delay model, different models have different assumptions made with different behaviors in both uncongested and congested situations. A research conducted by Dion et al. (2004) compared vehicle delays provided by a number of analytical delay models with delays estimated by microscopic traffic simulator on a one-lane approach to a pre-timed signalized intersection approach for traffic conditions ranging from under-saturation to over-saturation. Among all types of delay formulations, the time-dependent stochastic delay models were reported to provide strong consistency with the microscopic simulation method approach under both under-saturated and over-saturated conditions. As one of the most widely accepted time-dependent stochastic delay models, 2010 HCM delay model employed the incremental queue accumulation procedure to calculate the uniform delay instead of the first item of the Webster's formula (Strong & Rouphail, 2006). The new method removed the aforementioned assumptions to allow more accurate uniform delay estimation for progressed traffic movement, movements with multiple green periods, and movements with multiple saturation flow rates, such as protected-permitted turn movement, which were most ignored in the existing methods of combined traffic control and assignment problems. This is very important when dealing with real world problem with complex traffic signal settings, especially for urban networks.

2.3 Solution Methods

To Solve the MPEC model, GAMS/NLPEC solver is one of the few or maybe the only tool in the market to solve an MPEC model. It reformulates the complementarity constraints and makes the MPEC problem into sequence of general Nonlinear Programming (NLP) models which can be solved by existing NLP solvers in GAMS. Then, it extracts the MPEC solution from the NLP solution. The reformulated models NLPEC produces are in scalar form. Different reformulations methods are supported by the NLPEC solver and the combination of different reformulations and NLP solvers produces a good chance to solve the problem efficiently and effectively.

However, for the MINLP with Complementarity Constraints, NLPEC solver is not capable to solve the problem in that it is only able to deal with problem with continuous variables only. Heuristic algorithm will be used in this study. One of the most widely used Heuristic algorithm in signal optimization problem is the genetic algorithm (GA) (Yin, 2000; Ceylan & Bell, 2004; Ceylan & Bell, 2005; Ceylan, 2006). In this study, GA will be employed as a candidate algorithm to solve the proposed MINLP with Complementarity

Constraints.

The GA begins its iterative computation with a population of random strings representing the decision variables. The population is then operated by three main operators: reproduction (selection), crossover, and mutation to create a new population of points. The reproduction operator selects strings which are better than others and the crossover operator recombines good strings together to create a new better string, while the mutation operator alters a string locally expecting a better string. Basically, at each of these three steps it is expected that if a bad string has been produced it will be removed from the population and those with good features will be part of the new population. This new population will be used to generate the next population and at each step the fitness of the new generation can be obtained as the value of the objective function. In each generation if the solution is improved, it is stored as the best solution. The basic steps for the GA computation are as follows.

Table 2.1. Basic Steps of Genetic Algorithm

Algorithm
Generate Initial Population, *P*
Evaluate (P)
$a \leftarrow 0$
$X \leftarrow$ the best solution in P
while stopping condition not met **do**
$a \leftarrow a + 1$
Parent Selection (*P*)
P' = Crossover (*P*)
P'' = Mutation (*P'*)
P = Replacement (*P*, *P''*)
Evaluate (*P*)
if the best solution *Xa* in *P* is better than *X* **do**
return *Xa*
end if

end while

CHAPTER 3. THE PROPOSED BASIC MODEL

This chapter concentrates on the problem statements, proposed methodology, and model formulation of the basic MPEC model.

3.1 Problem Statement

This basic model attempts to formulate a mathematical problem that minimizes the travel cost, which includes both roadway segment travel time cost and signalized intersection control delay, in a given roadway transportation system by determining the optimal effective green times of the corresponding traffic signal system, while considers traffic flow equilibrium simultaneously.

Roadway segment travel time cost and signalized intersection control delay are the two components of the total system cost that are considered in the problem. h_a denotes the average vehicle travel time of link a, and it depends on link traffic flow f_a. $\boldsymbol{\Delta}$ is the link-path incidence matrix with elements $\delta_a^p = 1$, if path p traverse link a, 0 otherwise. Traffic signals are implemented on signalized intersections $n \in \boldsymbol{N}_s$. Cycle length of the signal is denoted by ω_n. Each signal has a given cycle which contains several phases $\boldsymbol{L}_n$ with corresponding effective green time for each phase denoted by $x_n^{l_n}$. d_{nai} is the average control delay per vehicle, which includes uniform delay d_{nai}^1 and incremental delay d_{nai}^2, for lane group nai. Therefore, total travel cost on path p per vehicle, c_p, can be obtained from the summation of travel time on each link and control delays at each signalized intersection traversed by this path.

In the proposed optimization model, urban network traffic signal timing optimization problem and user equilibrium traffic assignment problem are considered as a combined optimization problem. The combined problem is formulated as a mathematical model that attempts to minimize total travel time, $\sum_p f_p c_p$, with decision variables that are signal timing parameters, in particular green splits $x_n^{l_n}$. The user equilibrium traffic assignment is taken into account as a set of constraints in the formula. HCM 2010 delay method, which is one of the most up-to-date time-dependent stochastic delay models, is employed as intersection delay estimation method in the model. The proposed combined problem belongs to a class of challenging optimization problems, namely mathematical programming with equilibrium constraints (MPEC). The complementarity constraints are necessary to model the user equilibrium traffic assignment condition and delay constraints.

3.2 Mathematical Model

This section describes the mathematical model which is formulated to represent the proposed optimization problem introduced in the previous section.

3.2.1 Network Definition. Consider a traffic network represented by a directed network $G(\boldsymbol{N}, \boldsymbol{A})$, where $\boldsymbol{N}$ is the set of nodes and $\boldsymbol{A}$ is the set of links. Nodes in the directed network can be signalized intersections, origins/destinations (O/D) of trips, or both. Among all, $\boldsymbol{N}_s$ and $\boldsymbol{N}_{od}$ represent the subset of nodes, $\boldsymbol{N}$, that include all signalized intersection and all O/D nodes, which produce and attract trips in the traffic network, respectively. Set of origin-destination (O-D) pairs is denoted by $\boldsymbol{O}$. For each O-D pair $o \in \boldsymbol{O}$, there exists a demand α_o. Links, $\boldsymbol{A}$, in the directed network represent directed roadway segments that connect intersections and O/D with one or multiple traffic lanes. $\boldsymbol{A}_n$ is the subset of links that have a common head $n \in \boldsymbol{N}$. When approaching a signalized intersection $n \in \boldsymbol{N}_s$, travel lanes are categorized into left-turn lanes, right-turn lanes, or through lanes in terms of different traffic movements, left turn, right turn, and through. In traffic signal operation, traffic movements that do not conflict with each other are generally allowed to move at the same time, in the same signal phase, to be exact. Therefore, traffic lanes in all links that head to an signalized intersection are summarized into two lane groups, which are left-turn (LT) lane group, i_1, and through and right-turn (TRT) lane group, i_2. $\boldsymbol{I}$ denotes the set of these two lane groups. Then, in this problem, any specific lane group can be located by using the combination of n, a, and i. For instance, in Figure 3, northbound left-turn lane group of intersection n_1 is denoted as $n_1 a_1 i_1$ in the model.

Three types of traffic flows are defined in this problem, including path flow f_p, link flow f_a, and lane group flow f_{nai}. The set of possible paths across all O-D pairs is denoted by $\boldsymbol{P}$. $\boldsymbol{P}_a$ is the set of paths that traverse link a. The set of paths that go through a lane group nai is denoted by $\boldsymbol{P}_{nai}$. For each O-D pair $o \in \boldsymbol{O}$, there exists a travel demand α_o. Every O-D pair with non-zero demand will generate traffic flows on one or more paths that connect this O-D, and traffic flow on path p is denoted by a path flow f_p. One roadway segment a may be on multiple different paths. Link flow is denoted by f_a and can be obtained from path flows by $\sum_{p \in \boldsymbol{P}_a} f_p$. When approaching to a signalized intersection, traffic flows are summarized by lane groups $\boldsymbol{I}$. Lane group flow f_{nai} is also able to be obtained from path flows by $\sum_{p \in \boldsymbol{P}_{nai}} f_p$. Path flow f_p, link flow f_a, and lane group flow f_{nai} will be used in user equilibrium traffic assignment, link travel cost calculation, and control delay calculation, respectively.

All three types of flows are important variables in the proposed model and represent the traffic assignment in the system. Another variable that plays a similar role in the model is the effective green length for each phase, which represents the traffic signal settings. To address the problem, we define the decision variable, $x_n^{l_n}$, to be the length of effective green time of each phase, $l_n \in \boldsymbol{L}_n$ at any signalized intersection, $n \in \boldsymbol{N}_s$. The total travel cost on path p equals to the summation of travel time on each link, $\sum_a \delta_a^p h_a$, and control delays at each signalized intersection traversed by the path, $\sum_{n \in \boldsymbol{N}_s, a \in \boldsymbol{A}_n, i} \gamma_{nai}^p d_{nai}$. Then the objective function that attempt to minimize the system total travel cost is represented as

$$Min \sum_p f_p c_p$$

Subject to

$$c_p = \sum_a \delta_a^p h_a + \sum_{n \in N', a \in A_n, i} \gamma_{nai}^p d_{nai}, \forall\, p$$

The user equilibrium traffic assignment problem is simultaneously considered as a set of constraints and a set of classic complementarity constraints is adopted. In this model, τ_o denotes the minimum travel cost among all the paths that connect O-D pair o. $\boldsymbol{\Theta}$ is the O-D-path incidence matrix with elements $\theta_o^p = 1$, if path p connect O-D pair o, 0 otherwise. User equilibrium traffic assignment constraints are listed as below.

$$0 \leq f_p \perp c_p - \sum_o \theta_o^p \tau_o \geq 0, \forall\, p \in \boldsymbol{P}$$

$$\sum_p \theta_o^p f_p - \alpha_o = 0, \forall\, o \in \boldsymbol{O}$$

$$\tau_o \geq 0, \forall\, o \in \boldsymbol{O}$$

The first Equation of the three Equations shown above indicates that under user equilibrium traffic flow, any used path will have the same and minimum travel cost among all the paths, $\sum_o \theta_o^p \tau_o$, for any origin-destination pair (when $0 \leq f_p$, $c_p - \sum_o \theta_o^p \tau_o = 0$ has to be satisfied), otherwise no flow uses this path (when $0 = f_p$, $c_p - \sum_o \theta_o^p \tau_o \geq 0$ has to be satisfied). Meanwhile, The second Equation ensures that, for each O-D pair, the traffic flows satisfy the traffic demand α_o.

3.2.2 Link Travel Time Estimation. In the proposed model, calculation of link travel time uses the well-known BPR function showing as below, where h_a is the estimated link travel time, ff_a is the link free flow travel time, f_a is the link flow, and s_a is the link saturation flow (link capacity).

$$h_a = ff_a \cdot \left[1 + 0.15\left(\frac{f_a}{s_a}\right)^4\right]$$

Besides, a minimum and maximum green duration, $gmin_n^l, gmax_n^l$, for each phase are also taken into account. Meanwhile, for different phase that includes different types of lane group, the corresponding minimum and maximum green interval might be different. The minimum green duration represents the least amount of time that a green signal will be displayed for a moment, or a lane group. It is determined by the time drivers need to go through the intersection, pedestrian crossing time, etc. Normally, the minimum green duration for a through movement is in the range of 2 to 15 seconds, which depends on the facility type such as major arterial and minor arterial, while for left turn movement, the minimum green time needed is always shorter, 2 to 5 seconds. The maximum green interval is used to limit the delay to any other movement at the intersection and to keep the cycle length to a maximum amount. Similar to minimum green duration, the maximum green duration for a through movement varies from 20 to 70 seconds based on the facility type, while for left turn movement, 15 to 30 seconds.

3.2.3 Intersection Delay Calculation. As the main contribution of the proposed approach, the new control delay calculation method implemented in this model will be explained and discussed in this section. In this model, HCM 2010 method is employed to determine intersection signal control delays for a lane group. Two important reasons of choosing to use this method are 1) its capability and reliability to estimate the control delays under both under-saturated and over saturated situations and 2) its capability to provide more accurate uniform delay estimations for movements with multiple green periods and multiple control methods, such as permitted turning movements. In the original HCM 2010 method, signal control delay for a lane group is considered as the combination of three components, which are uniform delay, incremental delay, and initial queue delay. In this model, it is assumed that at the beginning of the analysis period, no initial queue exists for any lane group at any intersection. Under this assumption, the situation without initial queue delay is considered in this study.

3.2.3.1 Lane Group without Permitted Left Turn. According to HCM 2010 method, the calculation of uniform delay for a lane group is based on the area bounded the polygon shown in Figure 1, which is used for lane groups that do not have permitted left turn. The set of all lane groups that have no permitted left turn is denoted as $\boldsymbol{NAI}_1$.

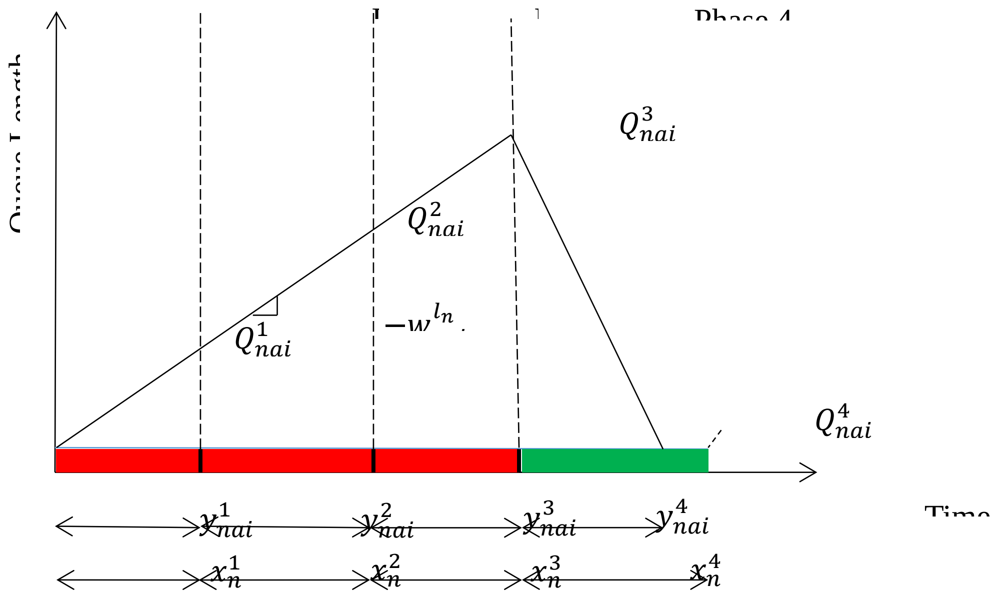


Figure 3.1. Uniform Delay Shape for Normal Lane Groups

Figure 3.1 is a polygon shape that illustrate the uniform delay for a through and right turn lane group with a 4-phase signal timing plan. $Q_{nai}^{l_n}$ is the queue length of this lane group at the end of phase l_n. $w_{nai}^{l_n}$ is the queue change rate and $y_{nai}^{l_n}$ is the queue change duration.

The area bounded by the polygon represents the total uniform delay, and then the total is divided by the number of arrivals per cycle to estimate the average uniform delay. Thus, in HCM 2010, these calculations are summarized in the equations below.

$$d_{nai}^1 = \frac{\sum_{l_n}\left[0.5 \cdot (Q_{nai}^{l_n-1} + Q_{nai}^{l_n}) \cdot y_{nai}^{l_n}\right]}{v_{nai}\omega_n}, \forall\, nai \in \boldsymbol{NAI}_1$$

$$Q_{nai}^{l_n} = Q_{nai}^{l_n-1} - y_{nai}^{l_n} \cdot w_{nai}^{l_n}, \forall\, nai \in \boldsymbol{NAI}_1$$

$$y_{nai}^{l_n} = \min(x_n^{l_n}, \frac{Q_{nai}^{l_n-1}}{w_{nai}^{l_n}}), \forall\, l_n, \forall\, nai \in \boldsymbol{NAI}_1$$

In which, the queue change (build-up or vanish) rate, $w_{nai}^{l_n} = s_{nai}^{l_n} - \frac{v_{nai}}{lnum_{nai}}$, and $Q_{nai}^{l_n-1}$ denotes the queue length at the end of l_n's previous phase. $lnum_{nai}$ is the number of lanes of this lane group. It has to be mentioned that Equation (1) is not mathematically rigorous. For the phases that queue length decreases, $w_{nai}^{l_n} > 0$, the value of the queue change duration $y_{nai}^{l_n}$ is able to be successfully determined by Equation (3) since the queue clearance time is nonnegative, $\frac{Q_{nai}^{l_n-1}}{w_{nai}^{l_n}} > 0$; however, when queue length does not decrease, such as phase 1, 2, or 3 in Figure 1, item $\frac{Q_{nai}^{l_n-1}}{w_{nai}^{l_n}}$ will be either negative or meaningless since queue change rate $w_{nai}^{l_n}$ is not positive in this case. The negative value of the queue clearance time, $\frac{Q_{nai}^{l_n-1}}{w_{nai}^{l_n}}$, will lead to a negative value of queue change duration, $y_{nai}^{l_n}$, due to the minimum logic, which contradicts with the fact that queue change duration should always be nonnegative. In addition, item $\frac{Q_{nai}^{l_n-1}}{w_{nai}^{l_n}}$ will become meaningless in mathematical model when queue change rate $w_{nai}^{l_n} = 0$. Thus, in order to keep the consistency of the formulation, the minimum logic in the original model has to be replaced by other mathematical logic that is more rigorous.

In this study, the following equations are proposed.

$$d_{nai}^{1} \cdot v_{nai} \cdot \omega_n = \sum_{l_n} \left[0.5 \cdot (Q_{nai}^{l_n-1} + Q_{nai}^{l_n}) \cdot y_{nai}^{l_n}\right], \forall\, nai \in \boldsymbol{NAI}_1$$

$$Q_{nai}^{l_n} = \max(0, Q_{nai}^{l_n-1} - x_n^{l_n} \cdot w_{nai}^{l_n}), \forall\, l_n, \forall\, nai \in \boldsymbol{NAI}_1$$

$$Q_{nai}^{l_n} = Q_{nai}^{l_n-1} - y_{nai}^{l_n} \cdot w_{nai}^{l_n}, \forall\, l_n, \forall\, nai \in \boldsymbol{NAI}_1$$

When observing the value of the queue length at the end of a phase, $Q_{nai}^{l_n}$, there are only two potential results: 1) 0, if queue fully cleared during this phase, and in this case, $Q_{nai}^{l_n} \geq Q_{nai}^{l_n-1} - x_n^{l_n} \cdot w_{nai}^{l_n}$, or 2) $Q_{nai}^{l_n-1} - x_n^{l_n} \cdot w_{nai}^{l_n}$, if queue is not able to be cleared during this phase, and in this case, it can be either shortening , $w_{nai}^{l_n} \geq 0$, or building the queue, $w_{nai}^{l_n} \leq 0$. Thus, with Equation (2), $Q_{nai}^{l_n}$ can be successfully obtained with any value of queue change rate $w_{nai}^{l_n}$. Queue change duration $y_{nai}^{l_n}$ is then constrained in Equation (3), and will be ranged between 0 and phase length $x_n^{l_n}$ internally. Therefore, the proposed equations cover all the possible situations in practice, and can be used to replace the minimum condition in the original formula.

The second equation can be modified into the simple complementarity conditions shown as follow.

$$0 \leq Q_{nai}^{l_n} \perp \left(Q_{nai}^{l_n} \geq Q_{nai}^{l_n-1} - x_n^{l_n} \cdot w_{nai}^{l_n}\right), \forall\, l_n, \forall\, nai \in \boldsymbol{NAI}_1$$

The above complementarity constraints hold only if at least one of the following holds.

$$0 \leq Q_{nai}^{l_n}, Q_{nai}^{l_n} = Q_{nai}^{l_n-1} - x_n^{l_n} \cdot w_{nai}^{l_n}$$

$$0 = Q_{nai}^{l_n}, Q_{nai}^{l_n} \geq Q_{nai}^{l_n-1} - x_n^{l_n} \cdot w_{nai}^{l_n}$$

To identify the previous phase of phase l_n, incidence matrix $\mathbf{Z}$, in which element $\zeta_{nai}^{l_n{}', l_n} = 1$, if phase $l_n{}'$ is the previous phase of l_n $(l_n{}', l_n \in \boldsymbol{L}_n)$ in one cycle, 0 otherwise, is introduced in the proposed model. The calculation of uniform delay is started from the first red light phase, and it is assumed that there is no initial queue, which means the queue length at the beginning of the first red light phase is 0. Thus, in order to achieve this, for the first red light phase of any lane group, all the elements are 0.

$$Q_{nai}^{l_n-1} = \sum_{l_n'} \zeta_{nai}^{l_n',l_n} \cdot Q_{nai}^{l_n'}$$

By using HCM 2010 method, the calculation of incremental delay can be formulated as

$$d_{nai}^2 = 900T\left[\left(\frac{v_{nai}}{\bar{s}_{nai}} - 1\right) + \sqrt{(\frac{v_{nai}}{\bar{s}_{nai}} - 1)^2 + \frac{4v_{nai}}{\bar{s}_{nai}{}^2 T}}\right], \forall\, nai \in \boldsymbol{NAI}_1$$

where

$$\bar{s}_{nai} = lnum_{nai} \times \sum_{l_n} (s_{nai}^{l_n} \times \frac{x_n^{l_n}}{\omega_n}), \forall\, nai \in \boldsymbol{NAI}_1$$

Since the capacity of lane group i, $\bar{s}_{nai}$, is a variable and always non-negative, in order to avoid numerical computation issues as $\bar{s}_{nai} = 0$, we modify the original incremental delay formulation to the following.

$$d_{nai}^2 \cdot \bar{s}_{nai} = 900T\left[(v_{nai} - \bar{s}_{nai}) + \sqrt{(v_{nai} - \bar{s}_{nai})^2 + \frac{4v_{nai}}{T}}\right], \forall\, i, \forall\, a \in \boldsymbol{A}_n, \forall\, n \in \boldsymbol{N}_s$$

3.2.3.2 Lane Group with Permitted Left Turn. Furthermore, Figure 3.2 illustrates the uniform delay for a left-turn lane group with permitted left turn movement allowed in phase 4. For this lane group, there is one protected phase, phase 3, and one permitted phase 4. The set of all lane groups that have permitted left-turn phase is denoted as NAI_2. In the permitted phase, left-turn traffic is allowed to make the left-turn maneuver once the opposing through and right-turn queue cleared. Thus, different from the case in Figure 3.1, delay polygon shape for the permitted phase, phase 4, is a combination of two trapezoids rather than one.

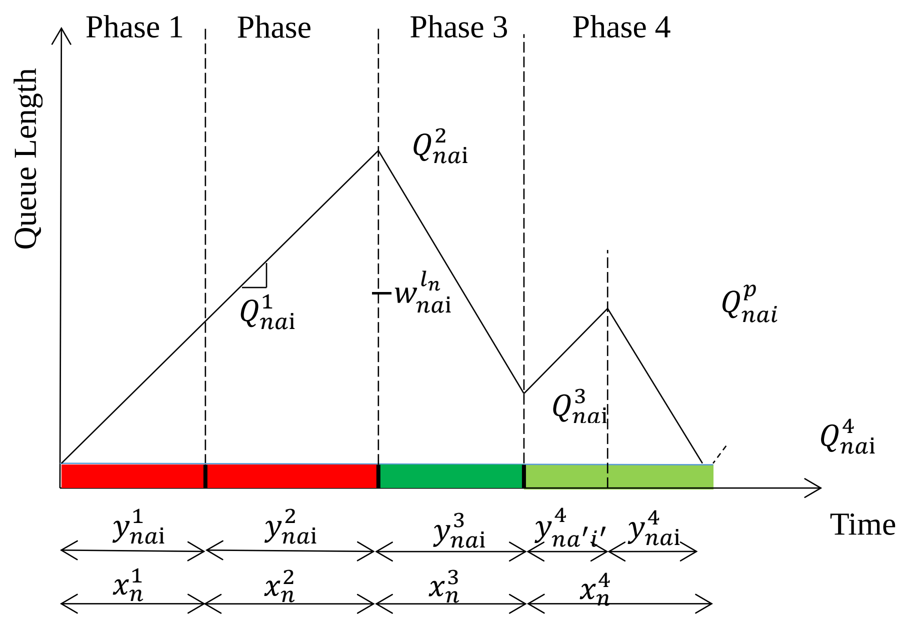


Figure 3.2. Uniform Delay Shape for Lead Phasing Left-Turn (LT) Lane Groups

In this case, same formula as provided in last section can be used for all other phases except the phase that has permitted left turn movement. Permitted phase is denoted as $l_n^p \in \boldsymbol{L}_n$, and in Figure 3.2, phase 4 is the permitted phase. For the permitted phase l_n^p, queue will firstly build-up until the oncoming TRT lane group queue being cleared, which needs time $y_{na'i'}^{l_n^p}$, after that, permitted left turn maneuver can start and vanishes the longer queue, Q_{nai}^p. Thus, uniform delay formula introduced in last section can be modified as follows for this case.

$$d_{nai}^1 \cdot v_{nai} \cdot \omega_n = \sum_{l_n \in (\boldsymbol{L}_n \setminus l_n^p)} \left[\frac{\left(Q_{nai}^{l_n-1} + Q_{nai}^{l_n}\right) \cdot y_{nai}^{l_n}}{2} \right] + \frac{\left(Q_{nai}^{l_n^p-1} + Q_{nai}^p\right) \cdot y_{na'i'}^{l_n^p} + \left(Q_{nai}^p + Q_{nai}^{l_n^p}\right) \cdot y_{nai}^{l_n^p}}{2}$$

$$\forall\, nai \in \boldsymbol{NAI}_2$$

$$0 \le Q_{nai}^{l_n} \perp \left(Q_{nai}^{l_n} \ge Q_{nai}^{l_n-1} - x_n^{l_n} \cdot w_{nai}^{l_n}\right), \forall\, l_n \in \left(\boldsymbol{L}_n \setminus l_n^p\right)$$

$$\forall\, nai \in \boldsymbol{NAI}_2$$

$$0 \le Q_{nai}^{l_n^p} \perp \left[Q_{nai}^{l_n^p} \ge Q_{nai}^{l_n^p-1} - \left(x_n^{l_n^p} - y_{na'i'}^{l_n^p}\right) \cdot w_{nai}^{l_n^p}\right]$$
$$\forall\, nai \in \boldsymbol{NAI}_2$$

$$Q_{nai}^{l_n} = Q_{nai}^{l_n-1} - y_{nai}^{l_n} \cdot w_{nai}^{l_n}, \forall\, l_n \in \left(\boldsymbol{L}_n \backslash l_n^p\right) \qquad \forall\, nai \in \boldsymbol{NAI}_2$$

$$Q_{nai}^{l_n^p} = Q_{nai}^{\mathrm{p}} - y_{nai}^{l_n^p} \cdot w_{nai}^{l_n} \qquad \forall\, nai \in \boldsymbol{NAI}_2$$

$$Q_{nai}^{\mathrm{p}} = Q_{nai}^{l_n^p-1} - y_{na'i'}^{l_n^p} \cdot w_{nai}^{\mathrm{p}} \qquad \forall\, nai \in \boldsymbol{NAI}_2$$

where l_n^p is the permitted phase, Q_{nai}^{p} is the maximum queue length in the permitted phase, $na'i'$ denotes the oncoming through and right turn lane group, and w_{nai}^{p} is the queue change rate before reaching Q_{nai}^{p}, which equals to $0 - \frac{v_{nai}}{ln_{nai}}$. For other phases, $l_n \in \left(\boldsymbol{L}_n \backslash l_n^p\right)$, calculation of uniform delay, as Equation (4)'s first item of right hand side, Equation (5), and Equation (7), is formulated in the same way as other lane groups.

While for the permitted phase, since queue length does not increase or decrease consistently, it is not proper to estimate the uniform delay by calculating area of one trapezoid using queue length at the beginning and end of the phase. Equation (9) is introduced to get maximum queue length Q_{nai}^{p}. With Q_{nai}^{p}, similar to Equation (5) and (7), Equation (6) and (8) can constraint the queue length at the end of this phase $Q_{nai}^{l_n^p}$. Then, the second item of the right hand side of Equation (4) calculates the total area of the two trapezoids in the permitted, which is also the average uniform delay of this permitted phase.

3.2.4 Model Formulation.

Complete model formulation is provided in this section followed by a brief introduction of objective function and all constraints. The proposed approach minimizes the total vehicle costs, which consist of both link travel time and intersection delays. Several factors are simultaneously considered, including user equilibrium traffic assignment, time-dependent stochastic intersection control delay, and their interactive impacts to each other. In addition, operational feasibility of signal settings such as the required minimum and maximum green time for each phase is also considered in the model. To avoid the difficulty following the formulation, a list of all notations used in this study is provided in List of Symbols at the beginning of the report.

$$in \sum_{p} f_p c_p$$

Subject to

$$0 \leq f_p \perp c_p - \sum_{o} \theta_o^p \tau_o \geq 0, \forall\, p \in \boldsymbol{P}$$

$$\sum_{p} \theta_o^p f_p - \alpha_o = 0, \forall\, o \in \boldsymbol{O}$$

$$\sum_{l_n} \left(x_n^{l_n} + t_{lost}\right) = \omega_n, \forall\, n \in \boldsymbol{N}_s$$

$$c_p = \sum_{a} \delta_a^p h_a + \sum_{n \in \boldsymbol{N}_s, a \in \boldsymbol{A}_n, i} \gamma_{nai}^p d_{nai}, \forall\, p$$

$$d_{nai} = d_{nai}^1 + d_{nai}^2, \forall\, i, \forall\, a \in \boldsymbol{A}_n, \forall\, n \in \boldsymbol{N}_s$$

$$f_{nai} = \sum_{p \in \boldsymbol{P}_{nai}} f_p, \forall\, i, \forall\, a \in \boldsymbol{A}_n, \forall\, n \in \boldsymbol{N}_s$$

$$d_{nai}^{1} \cdot v_{nai} \cdot \omega_n = \sum_{l_n} \left[\frac{\left(\sum_{l_n'} \zeta_{nai}^{l_n',l_n} \cdot Q_{nai}^{l_n'} + Q_{nai}^{l_n}\right) \cdot y_{nai}^{l_n}}{2} \right], \forall\, nai \in \boldsymbol{NAI}_1$$

$$d_{nai}^{1} \cdot v_{nai} \cdot \omega_n = \sum_{l_n \in (\boldsymbol{L}_n \setminus l_n^p)} \left[\frac{\left(Q_{nai}^{l_n-1} + Q_{nai}^{l_n}\right) \cdot y_{nai}^{l_n}}{2} \right] + \frac{\left(Q_{nai}^{l_n^p-1} + Q_{nai}^{p}\right) \cdot y_{na'i'}^{l_n^p} + \left(Q_{nai}^{p} + Q_{nai}^{l_n^p}\right) \cdot y_{nai}^{l_n^p}}{2}, \forall\, nai \in \boldsymbol{NAI}_2$$

$$d_{nai}^{2} \cdot \bar{s}_{nai} = 900T \left[(v_{nai} - \bar{s}_{nai}) + \sqrt{(v_{nai} - \bar{s}_{nai})^2 + \frac{4v_{nai}}{T}} \right],$$

$$\forall\, i, \forall\, a \in \boldsymbol{A}_n, \forall\, n \in \boldsymbol{N}_s$$

$$Q_{nai}^{l_n} = \sum_{l_n'} \zeta_{nai}^{l_n',l_n} \cdot Q_{nai}^{l_n'} - y_{nai}^{l_n} \cdot w_{nai}^{l_n}, \forall\, nai \in \boldsymbol{NAI}_1$$

$$Q_{nai}^{l_n} = \sum_{l_n'} \zeta_{nai}^{l_n',l_n} \cdot Q_{nai}^{l_n'} - y_{nai}^{l_n} \cdot w_{nai}^{l_n}, \forall\, l_n \in \left(\boldsymbol{L}_n \setminus l_n^p\right), \forall\, nai \in \boldsymbol{NAI}_2$$

$$Q_{nai}^{l_n^p} = Q_{nai}^{\mathrm{p}} - y_{nai}^{l_n^p} \cdot w_{nai}^{l_n}, \forall\, nai \in \boldsymbol{NAI}_2$$

$$Q_{nai}^{\mathrm{p}} = \sum_{l_n'} \zeta_{nai}^{l_n',l_n^p} \cdot Q_{nai}^{l_n'} - y_{na'i'}^{l_n^p} \cdot w_{nai}^{\mathrm{p}}, \forall\, nai \in \boldsymbol{NAI}_2$$

$$w_{nai}^{l_n} = s_{nai}^{l_n} - \frac{v_{nai}}{lnum_{nai}}, \forall\, l, \forall\, i, \forall\, a \in \boldsymbol{A}_n, \forall\, n \in \boldsymbol{N}_s$$

$$\bar{s}_{nai} = lnum_{nai} \times \sum_{l_n} \left(s_{nai}^{l_n} \times \frac{x_n^{l_n}}{\omega_n}\right), \forall\, i, \forall\, a \in \boldsymbol{A}_n, \forall\, n \in \boldsymbol{N}_s$$

$$0 \le Q_{nai}^{l_n} \perp \left(Q_{nai}^{l_n} \ge \sum_{l_n'} \zeta_{nai}^{l_n',l_n} \cdot Q_{nai}^{l_n'} - x_n^{l_n} \cdot w_{nai}^{l_n}\right),\ \forall\, l_n, \forall\, nai \in \boldsymbol{NAI}_1$$

$$0 \le Q_{nai}^{l_n} \perp \left(Q_{nai}^{l_n} \ge \sum_{l_n'} \zeta_{nai}^{l_n',l_n} \cdot Q_{nai}^{l_n'} - x_n^{l_n} \cdot w_{nai}^{l_n}\right),$$

$$\forall\, l_n \in \left(\boldsymbol{L}_n \backslash l_n^p\right), \forall\, nai \in \boldsymbol{NAI}_2$$

$$0 \le Q_{nai}^{l_n^p} \perp \left[Q_{nai}^{l_n^p} \ge \sum_{l_n{}'} \zeta_{nai}^{l_n{}', l_n^p} \cdot Q_{nai}^{l_n{}'} - \left(x_n^{l_n^p} - y_{na'i'}^{l_n^p}\right) \cdot w_{nai}^{l_n^p}\right],$$

$$\forall\, nai \in \boldsymbol{NAI}_2$$

$$\boldsymbol{gmin_n^{l_n}} \le \boldsymbol{x_n^{l_n}} \le \boldsymbol{gmax_n^{l_n}}, \forall\, \boldsymbol{l}, \forall\, \boldsymbol{n} \in \boldsymbol{N'}$$

$$\boldsymbol{f_p,\ h_a,\ c_p,\ Q_{nai}^{l_n},\ Q_{nai}^{\mathbf{p}},\ v_{nai},\ \bar{s}_{nai},\ y_{nai}^{l_n},\ d_{nai},\ d_{nai}^{1}}, \textbf{and}\ \boldsymbol{d_{nai}^{2}} \ge \mathbf{0}$$

The objective function (10) minimizes the total travel cost in the roadway transportation network. The constraints as shown in Equation (11) and (12) ensure that user equilibrium traffic assignment is always satisfied. Equation (13) ensures the feasibility of signal timing settings. Equation (14) gives the total travel cost per vehicle of one path. While, the combination of constraints expressed by Equation (15) to (21) constraint the control delays of every lane group in the network, including both with and without permitted left-turn phase. Equation (15) ensures that control delay include both uniform delay and incremental delay. Equation (16) ensures that lane group flow is obtained from the corresponding path flows. Equation (17), (19), (20), (21), and (22) constraint the uniform delay for both lane groups with permitted left-turn phase or without permitted left-turn phase by using corresponding formula. Equation (18) constraints the incremental delay for all lane groups. In addition, Equations (23) bound the value of green splits in the pre-defined min-max range. Finally, Equation (24) ensures the feasibility of all the non-negative variables in the model.

CHAPTER 4. COMPUTATIONAL EXPERIMENTS USING THE BASIC MODEL

To solve the basic model, GAMS/NLPEC solver is adopted as it is one of the most accepted solvers that is able to handle a MPEC problem like the one described in this study. Proposed methodology is firstly implemented on a small network to verify the feasibility of the model application. Exhaustive interpretation of model settings and experiment procedures will be discussed to give a clear explanation on the general process of model application. Then, a larger network from the City of Chicago with field traffic demand data and signal settings will be used as test bed for verification of the practicability of model's real world application.

4.1 Computational Experiment I

4.1.1 Test Network and Settings. The small test network includes 4 intersections, $[n_1, n_2, n_3, n_4]$, with pre-timed signal control and 8 origin/destination nodes around, $[o_1, o_2, \dots, o_8]$. Nodes in the network are connected by 24 directed links, $[a_1, a_2, \dots, a_{24}]$, each of which represents one direction of the roadway segment that connects intersections and/or origin/destination. To simplify the input setting process, free flow travel time of any link in the network are assumed to be 15 seconds. All links that go into signalized intersections are assumed to have two lane groups of traffic, include a left-turn lane group i_1 and a through and right-turn lane group i_2. Figure 4.1 illustrates the small network described above.

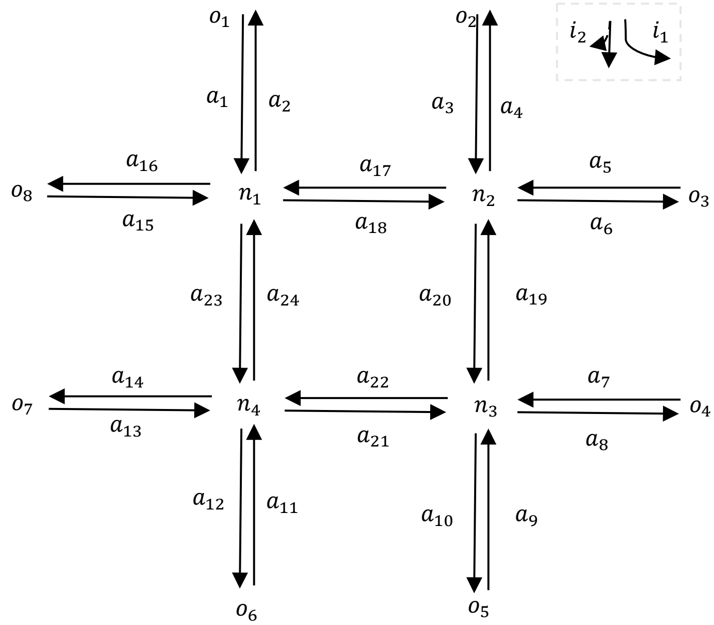


Figure 4.1. Network One

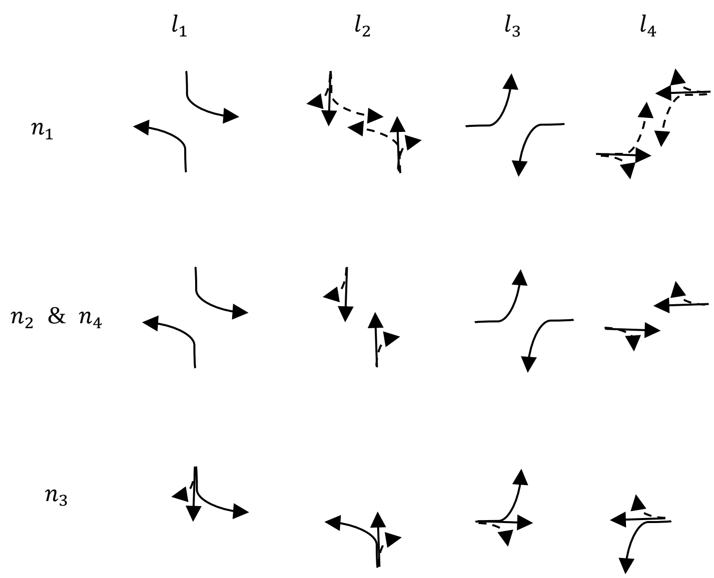


Figure 4.2. Phase Plans for Small Network

In order to verify the proposed model’s capability of modelling different types of intersection signal control, 3 different pre-timed phasing plans are implemented. At intersection n_1, a default 4-phase protected-permitted left-turn phasing is used, which includes a combination of protected left-turn phase, which only allows left-turns move, and permitted left-turn phase, which allows left-turn moves after yielding to conflicting traffic and pedestrians. While for intersection n_2 and n_4, protected-only left-turn phasing is employed as no permitted left-turn movement is allowed. For these three intersections, east west (EW) left-turns move in phase l_1, which is followed by moving of EW through movements in phase l_2. Similar process happens for north south (NS) approaches in phase l_3 and l_4. In contrast with the phasing plans used at other 3 intersections, a split phasing plan is employed at intersection n_3. In the split phasing plan, all movements from one approach are allowed to move simultaneously in one phase, in which no movement from other approaches is allowed to move. Figure 4 illustrates the intersection phase settings used in this small network case. Besides, in the proposed model, all signal phases share one common min-man range as the effective green time should not be less than 1 seconds or more than 50 seconds.

For each OD pair, two potential paths are manually pre-selected based on two rules: 1) with least number of links and 2) with least number of left-turns. Obeying first rule will lead to findings of first and second shortest paths between this OD pair without the consideration of intersection control delays. While the second rule is based on an assumption that it is highly likely that left-turns create more delay than right-turns. Therefore, a total of 112 paths between 56 different OD pairs (exclude 8 OD pairs in which origin and destination are the same node) are considered in this small network case.

4.1.2 Initial Values and Bounds. GAMS/NLPEC is able to solve MPEC problems by converting complementarity constraints into general nonlinear constraints, and provide a local optimal solution. However, the solver may not able to generate a feasible solution without a proper setting of initial values and lower and upper bounds for certain variables.

Especially for the case when 0 is a valid value for most of the variables used in the model, such as the current study. Then, with a value equals to or very close to 0, it may cause a result of failing to find any feasible solution and terminate the optimization. Thus, choosing proper initial values and bounds for important variables are critical.

In order to obtain the initial values of effective green time, path flow, and lane group flow that are the closest to the optimal results, these three were pre-generated by using other model and tool.

Initial values of path flow f_p and lane group flow f_{nai} were determined by solving the proposed model without the constraints related to intersection control delays. Then, the problem became to a simple User Equilibrium Traffic Assignment problem which considers link travel time only.

With the initial values of lane group flow f_{nai}, initial effective green time $x_n^{l_n}$ can be then defined by maintain the same ratio with critical lane group volume, which is the general step to determine the optimal green splits in many signal optimization tools. In this study, Synchro version 8 is adopted as the tool to generate initial effective green time $x_n^{l_n}$. In addition, Synchro is also able to be used as a tool to measure the improvement of the traffic in network after implementing the proposed model.

4.1.3 Test Method. Three different traffic conditions, low demand, med demand, and high demand, are tested in the small network case with average V/C ratios ($\frac{v_{nai}}{\bar{s}_{nai}}$ in the model) for intersections range from 30% (low) to 150% (high). Traffic demands between OD pairs are randomly generated in order to keep the generality of the test results. By this test method, performance of the proposed model under different traffic loads were tested. OD Demands for all three demand conditions are listed in Table 4.1, 4.2, and 4.3.

Table 4.1. Low Demands (V/C = 30%)

veh/hr	o_1	o_2	o_3	o_4	o_5	o_6	o_7	o_8
o_1	0	82.8	25.2	46.8	18	28.8	46.8	18
o_2	61.2	0	61.2	50.4	3.6	25.2	50.4	3.6
o_3	25.2	32.4	0	50.4	7.2	50.4	18	18
o_4	61.2	64.8	64.8	0	54	54	46.8	64.8
o_5	43.2	46.8	43.2	43.2	0	10.8	43.2	25.2
o_6	7.2	32.4	3.6	10.8	21.6	0	10.8	32.4
o_7	36	108	111.6	93.6	64.8	18	0	28.8
o_8	36	7.2	57.6	3.6	3.6	43.2	32.4	0

Table 4.2. Medium Demands (V/C = 90%)

veh/hr	o_1	o_2	o_3	o_4	o_5	o_6	o_7	o_8
o_1	0	165.6	50.4	90	36	57.6	93.6	36
o_2	122.4	0	122.4	100.8	7.2	50.4	100.8	3.6
o_3	50.4	64.8	0	100.8	14.4	100.8	36	18
o_4	122.4	129.6	129.6	0	108	104.4	90	126
o_5	82.8	126	86.4	86.4	0	25.2	86.4	54
o_6	14.4	64.8	3.6	21.6	43.2	0	21.6	64.8
o_7	36	108	111.6	93.6	126	61.2	0	118.8
o_8	72	18	126	3.6	3.6	86.4	64.8	0

Table 4.3. High Demands (V/C = 150%)

veh/hr	o_1	o_2	o_3	o_4	o_5	o_6	o_7	o_8
o_1	0	252	79.2	133.2	57.6	86.4	140.4	57.6
o_2	180	0	183.6	151.2	7.2	75.6	151.2	3.6
o_3	75.6	93.6	0	147.6	21.6	154.8	57.6	28.8
o_4	180	194.4	190.8	0	162	158.4	136.8	187.2
o_5	126	190.8	133.2	133.2	0	36	133.2	82.8
o_6	21.6	97.2	7.2	36	64.8	0	32.4	97.2
o_7	57.6	162	169.2	140.4	190.8	90	0	176.4
o_8	108	28.8	187.2	7.2	7.2	129.6	100.8	0

4.1.4 Results and Findings. Results of model application under three different traffic conditions are presented in the following tables. Table 4.4, 4.5, and 4.6 list the hourly traffic flow loaded on each lane group before and after the optimization by using the proposed model, and signal timings are presented in Table 4.7 to 4.9. Meanwhile, detailed comparison of the optimization results from Synchro with initial traffic assignment and the proposed model is provided in Table 4.10 and Table 4.11.

Table 4.4. Low Demand Traffic Assignment Results (V/C = 30%)

f_{nai} (veh/hr)	EB		WB		NB		SB	
Intersection (Before)	LT	T&RT	LT	T&RT	LT	T&RT	LT	T&RT
1	36	148	68	209	112	266	115	151
2	169	194	58	144	101	295	61	194
3	76	256	54	356	122	133	101	94
4	212	248	90	234	11	108	65	220
(After)								
1	36	148	76	112	212	83	173	94
2	90	151	76	126	4	475	61	194
3	256	194	54	356	119	137	148	101
4	65	396	65	349	11	108	4	227

Table 4.5. Medium Demand Traffic Assignment Results (V/C = 90%)

f_{nai} (veh/hr)	EB		WB		NB		SB	
Intersection (Before)	LT	T&RT	LT	T&RT	LT	T&RT	LT	T&RT
1	72	302	112	436	328	338	216	313
2	356	292	198	187	241	472	122	385
3	0	418	108	702	112	436	202	198
4	374	281	223	403	22	212	133	414
(After)								
1	72	302	208	292	388	236	258	271
2	239	269	163	222	98	700	122	385
3	229	329	108	702	248	299	208	185
4	155	500	142	532	22	212	88	510

Table 4.6. High Demand Traffic Assignment Results (V/C = 150%)

f_{nai} (veh/hr)	EB		WB		NB		SB	
Intersection (Before)	LT	T&RT	LT	T&RT	LT	T&RT	LT	T&RT
1	108	461	212	590	421	565	342	464
2	464	454	169	410	302	785	184	569
3	76	626	162	1048	378	457	310	256
4	565	421	270	695	32	324	194	670
(After)								
1	108	461	132	483	528	362	481	325
2	378	490	249	330	195	988	184	569
3	279	473	162	1048	378	457	399	400
4	286	700	317	835	32	324	41	590

Table 4.7. Low Demand Signal Timing Results (V/C = 30%)

Intersection (Before) Synchro	Phase 1 (s)	Phase 2 (s)	Phase 3 (s)	Phase 4 (s)	Lost Time per Phase (s)	Cycle Length (s)
1	9	32	7	26	4	90
2	11	28	17	18		
3	11	14	21	28		
4	8	24	20	22		
(After)						
1	26.3	10.8	9.1	27.8	4	90
2	7.4	40.6	10.8	15.3		
3	13.5	12.5	22.4	25.6		
4	1.9	21.9	9.2	40.9		

Table 4.8. Medium Demand Signal Timing Results (V/C = 90%)

Intersection (Before) Synchro	Phase 1 (s)	Phase 2 (s)	Phase 3 (s)	Phase 4 (s)	Lost Time per Phase (s)	Cycle Length (s)
1	16	26	3	29	4	90
2	13	27	19	15		
3	8	21	17	28		

4	8	31	17	18		
(After)						
1	24.1	17.0	13.4	19.5	4	90
2	8.0	36.4	15.0	14.6		
3	10.9	14.7	15.4	33.0		
4	6.0	27.7	10.9	29.4		

Table 4.9. High Demand Signal Timing Results (V/C = 150%)

Intersection (Before) Synchro	Phase 1 (s)	Phase 2 (s)	Phase 3 (s)	Phase 4 (s)	Lost Time per Phase (s)	Cycle Length (s)
1	14	29	5	26	4	90
2	11	30	17	16		
3	10	16	18	30		
4	7	24	20	23		
(After)						
1	27.3	16.8	6.6	23.2	4	90
2	9.2	33.2	14.5	17.2		
3	15.2	16.1	14.7	27.9		
4	1.9	21.8	15.3	35.0		

Table 4.10. Results Comparison

Demand	Int	Avg. Delay per vehicle (s)		Delta in Avg. Delay per vehicle (%)	Delta in Intersection Total Delay (%)	Delta in Network Total Delay (%)	Delta in Network Total Cost (%)
		Synchro	Prop. Model				
Low	1	27.8	31.3	12.59	-4.83	-3.55	-2.55
	2	36.8	35.3	-4.08	-7.15		
	3	47	47.9	1.91	16.71		
	4	36.8	26.8	-27.17	-24.91		
Medium	1	40.1	52.4	30.67	25.12	-16.80	-14.54
	2	85.3	81	-5.04	-7.36		
	3	209.3	131.6	-37.12	-33.31		
	4	75.9	70.2	-7.51	-3.07		

High	1	130.8	104.2	-20.34	-27.46	-11.42	-10.80
	2	244.2	240.6	-1.47	-0.12		
	3	403.1	366.1	-9.18	-1.42		
	4	256.3	179.4	-30.00	-31.02		

Table 4.11. Control Delay and Link Travel Cost

Demand	Network Total Control Delay (min)		Network Total Link Travel Cost (min)		Delta in Network Total Control Delay (%)	Delta in Network Total Link Travel Cost (%)
	Synchro	Prop. Model	Synchro	Prop. Model		
Low	2920.17	2816.60	1175.42	1174.52	-3.55	-0.08
Medium	14816.92	12328.13	2152.18	2173.81	-16.80	1.00
High	56280.22	49852.77	3248.45	3248.71	-11.42	0.01

In all three cases, comparing with the Synchro optimization results under initial traffic assignment, the proposed model did improve system traffic condition in terms of both reducing the total delay in the network and reducing the total travel cost which consist of control delay and link travel cost in the network. While the improvements differ under various traffic conditions. When demand is low, system performance before and after optimization using the proposed model do not have large imparity as improvement of only 3.55% in delay reduction and 2.55% in total travel cost reduction, which is justified. Initial values of variables, especially the pre-generated ones, were already optimal based on only link travel time. And in the case of low demand, control delay is expected to be low and there is not much space for improvement.

The proposed model provides the best improvement in the case with medium demand (16.80% in delay reduction and 14.54% in total travel cost reduction) and a good improvement with high demand (11.42% in delay reduction and 10.81% in total travel cost reduction). These indicate that the proposed methodology will be highly profitable under near capacity traffic condition and able to provide reasonable benefit under over saturation traffic condition. One thing needs to be mentioned is that by using BPR method, increment of total link travel time is much less sensitive than increment of total control delay when traffic load increases. Besides, after implementation of the proposed model, total link travel time may increase or decrease. As shown in Table 4.11, changes in network total link travel cost before and after using the proposed model did not exceed 1 percent.

More findings can be obtained under a closer look on the new traffic assignment after optimization, as shown in Table 4.4 to 4.6. Phasing plan at intersection n_1 is almost the same as phasing plans at n_2 and n_3 except permitted left-turns are only allowed at intersection n_1. With this difference, it can be easily observed that after optimization, left-turn traffic at intersection n_1 are always increased (light red cell highlighted in tables). However, at the same time, left-turn traffic at intersection n_2 and n_4 are normally decreased after optimization (light blue cell highlighted in tables). The reason of this trend is that with permitted left-turn phasing, an intersections' capability to handle left-turns is better than protected left-turn phasing. But this advantage is weakened when demand becomes so high that oncoming traffic is endless. This is why under high demand case, this trend is not as significant as it appears in low and medium demand cases.

Besides, for intersection n_3, which uses split phasing plan, traffic flow tend to be equal for the two lane groups, left-turn and though and right-turn, in the same approach (light green cells highlighted in tables). This change will try to optimally utilizes the capacity of intersection n_3, which is able to reduce volume at other intersections while maintaining same performance at intersection n_3. Westbound is the only approach at intersection n_3 that always maintain the same traffic flow, and the reason of this is that since the network is small, traffic entering the network from westbound had limited choices.

All findings mentioned above strongly support the successful validation of the proposed model's feasibility and capability in model application.

4.2 Computational Experiment II

4.2.1 Real World Network and Settings. Figure 4.3 illustrates a real world network of a subarea in the City of Chicago's central district for the second computation experiment using the proposed basic model. This roadway network includes 13 signalized intersections (squares) and a total of 54 directed links. Intersection of E 33st Street and S La Salle Street is located at the south-east corner of network as node n_{13}, and at the north-east corner is the intersection of E 31st Street and S Martin Luther King Drive as node n_1. For traffic entering and leaving the network, 14 nodes are located around the network as origin/destination only nodes (circles). Unlike with the assumption made in the simple sample network, signalized intersections are also treated as the origin/destination for trips that start/end at those intersection or roadways entering those intersections in this network. This is more reality-oriented than only considering access nodes as origin/destination.

Traffic demands are generated by a well calibrated and validated microscopic simulation tool, Chicago TRANSIMS Model, which is able to provide simulation results of each and every travelers' trajectory and other information within a 24-hour period.

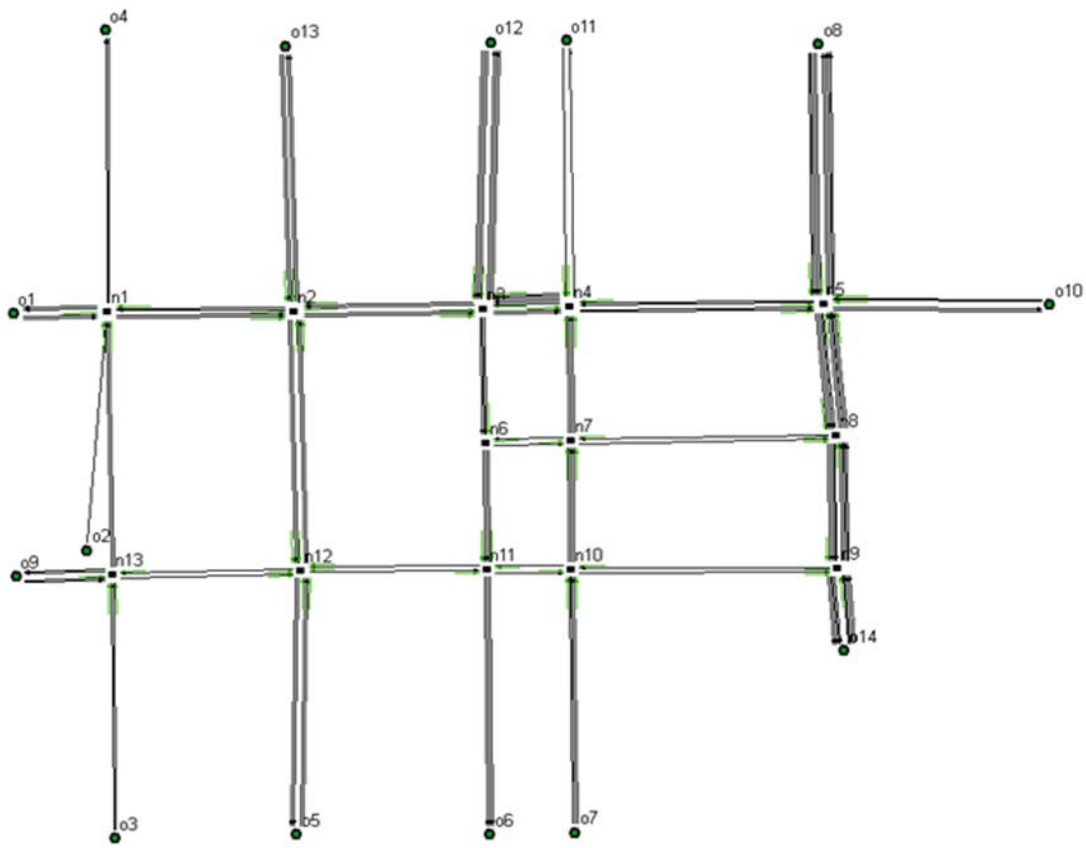


Figure 4.3. Network Two

Original signal timing plans coded in the model are based on the actual signal settings. 13 signalized intersections in the network belong to two different signal zones. Intersections along E 31st Street are parts of one signal zone with a cycle time equals to 85 seconds, while the other intersections belong to another signal zone with a 75-second cycle time. For all 13 signalized intersections, signal phases share one common min-man range as the effective green time should not be less than 1 seconds or more than 50 seconds.

For each OD pair in the network, up to 2 different candidate paths are pre-selected based on link free flow travel time by using the K shortest Loop-less Paths algorithm developed by Yen in 1971. All characteristics of the roadway network used in this case are either historical measurements or simulation results. In Table 4.12 below, important statistics of the network are listed. Number of candidate paths for different OD pairs are not necessarily the same, in that there may not exist more than one or two different loop-less paths for certain ODs. For some ODs, due to the existence of one way roads in the network, there are no paths and traffic demands.

Table 4.12. Network Statistics

Node		Candidate Paths	Number of Links	Link Length (ft)		Free Flow Travel Speed (MPH)	
Signalized	13	1117	54	Min	402	Min	26.6
OD Only	14			Max	1421	Max	43.4

__4.2.2 Test Method.__ Similar to the test method used in the previous case, Synchro version 8 is adopted as the tool to measure the potential improvement of the traffic performance in terms of total intersection delay caused by implementing the proposed model. Results of the proposed model are compared with the optimization results generated by using Splits Optimization in Synchro. The reason of not directly using the network before optimization as the reference of result comparison is the mismatching of signal timing plans and traffic assignment. The original traffic assignment is not from real field traffic count but based on a preliminary user equilibrium traffic assignment step in which only link travel time is considered, while the signal timing plans of the intersections are based on the real settings. Therefore, it is expected to observe a significant improvement on total intersection delay after optimization, either by using Synchro or the proposed method.

Besides, different traffic conditions are also tested in this real world case. Unlike the 4-intersection fictional network case, two different traffic conditions, peak and off-peak, are selected and separately used as input demand in the study. By using this test method, performance and capability of the proposed model under different traffic loads, including relatively dense traffic in peak hour in which congestion problems normally exist and also sparse traffic in off peak, are able to be tested in real world case.

4.2.3 Results and Findings. Results of model application on the real world network under peak hour demand condition is presented in Table 4.13 and 4.14, while results under off-peak hour demand condition (sparse demand) are listed in Table 4.15 and 4.16. Table 4.17 and 4.18 present the detailed comparison of the optimization results from Synchro with initial traffic assignment and the optimization results of the proposed model.

Table 4.13. Peak Hour Demand Traffic Assignment Results (AM peak)

f_{nai} (veh/hr)	EB		WB		NB		SB	
Int. # (Synchro)	LT	T&RT	LT	T&RT	LT	T&RT	LT	T&RT
1	107	631		397	40	271		
2	8	924	27	583	149	637	107	177
3	121	908	308	749			107	270
4	143	521		742	239	227	87	171
5	308	333	134	207	296	252	9	287
6			170				196	440
7	140	101		101	44	434		
8	63	67			26	420		326
9	0	153			69	302		288
10	255	289		127	114	223		
11		774	42	292			68	518
12	311	280	96	463	30	561	436	231
13	20	191		254	15	245		
(Proposed Model)								
1	107	631		366	40	359		
2	41	971	89	545	114	265	128	156
3	121	1038	206	736			95	282
4	65	508		479	363	390	87	171
5	212	488	142	199	27	286	9	287
6			111				196	701
7	98	109		146	110	725		
8	108	50			21	228		451
9	61	106			155	216		394
10	481	274		383	104	233		

11		697	42	402			150	635
12	56	478	73	542	30	561	19	258
13	17	193		332	15	245		

Table 4.14. Peak Hour Demand Signal Timing Results (AM peak)

Intersection (Before) Synchro	Phase 1 (s)	Phase 2 (s)	Phase 3 (s)	Phase 4 (s)	Lost Time per Phase (s)	Cycle Length (s)
1	17	23	28	1	4	85
2	9	24	3	33		85
3	13	25	35			85
4	10	21	11	27		85
5	9	19	19	22		85
6	24	43				75
7	23	21	19			75
8	21	20	22			75
9	26	18	19			75
10	22	22	19			75
11	3	48	12			75
12	9	21	15	14		75
13	7	17	19			75
(After)						
1	6.8	13.1	40.2	8.9	4	85
2	11.4	9.3	8.2	40.1		85
3	9.7	16.3	47			85
4	9.9	21.6	7.5	10		85
5	17.1	25.7	21.9	4.3		85
6	10.7	56.3				75
7	33.5	10.6	18.9			75
8	17.1	3.5	42.4			75
9	18.7	13.9	30.4			75
10	30.3	21.1	11.6			75
11	3.8	38.1	21.1			75
12	6.9	29.1	3.1	19.9		75
13	3.1	41.1	18.8			75

Table 4.15. Sparse Demand Traffic Assignment Results (Off-peak)

f_{nai} (veh/hr)	EB		WB		NB		SB	
Int. # (Synchro & Proposed Model)	LT	T&RT	LT	T&RT	LT	T&RT	LT	T&RT
1	5	15		7	0	1		
2	0	13	0	14	0	1	0	2
3	0	12	2	8			4	12
4	0	7		8	2	9	1	2
5	4	10	4	7	0	7	0	7
6			0				1	14
7	1	0		0	0	11		
8	0	2			0	7		11
9	0	2			0	7		13
10	6	1		3	1	9		
11		7	2	2			1	13
12	0	0	0	2	1	10	0	1
13	0	0		3	1	1		

Table 4.16. Sparse Demand Signal Timing Results (Off-peak)

Intersection (Before) Synchro	Phase 1 (s)	Phase 2 (s)	Phase 3 (s)	Phase 4 (s)	Lost Time per Phase (s)	Cycle Length (s)
1	5	25	31	8	4	85
2	1	31	6	31		85
3	5	42	26			85
4	5	7	18	39		85
5	25	27	5	12		85
6	5	62				75
7	5	47	11			75
8	5	1	57			75
9	5	1	57			75
10	1	41	21			75
11	1	37	25			75
12	1	14	7	37		75
13	1	27	35			75
(After)						

1	35.8	1	30.4	1.8	4	85
2	1	1	63.7	3.3		85
3	2.2	2.8	68.0			85
4	64.0	1.5	2.5	1		85
5	1.5	2.2	64.3	1		85
6	1	66				75
7	1	61	1			75
8	2.8	1	59.2			75
9	31.0	1	31.0			75
10	1	1.0	61.0			75
11	1	1.7	60.3			75
12	1.5	5.6	3.5	48.5		75
13	56.2	5.8	1			75

Table 4.17. Peak Hour Case Results Comparison

Case	Sig. Int.	Avg. Delay per vehicle (s)		Delta in Avg. Delay per vehicle (%)	Delta in Intersection Total Delay (%)	Delta in Network Total Delay (%)	Delta in Network Total Cost (%)
		Synchro	Prop. Model				
AM Peak	1	11.2	11.8	5.36	9.51	-20.53	-11.23
	2	26.6	26.8	0.75	-10.94		
	3	16.8	13.8	-17.86	-17.36		
	4	26.1	27.4	4.98	1.68		
	5	30.9	27.2	-11.97	-20.46		
	6	13.3	7.4	-44.36	-30.42		
	7	16.5	11.7	-29.09	2.73		
	8	10.6	11.8	11.32	5.89		
	9	16.7	8	-52.10	-45.02		
	10	17.6	22.2	26.14	84.57		
	11	18.3	19.3	5.46	19.91		
	12	57	24.3	-57.37	-64.29		
	13	12.2	11.7	-4.10	6.09		
Off Peak	1	6.9	2.5	-63.77	-63.77	-36.81	
	2	17.2	8.8	-48.84	-48.84		
	3	15.5	11.1	-28.39	-28.39		

4	17.7	13.6	-23.16	-23.16
5	23.9	17.6	-26.36	-26.36
6	0.6	0.6	0.00	0.00
7	4.9	4.6	-6.12	-6.12
8	14.2	2	-85.92	-85.92
9	9.8	2	-79.59	-79.59
10	15.5	16.7	7.74	7.74
11	18	13.3	-26.11	-26.11
12	19.6	5.8	-70.41	-70.41
13	11	9.6	-12.73	-12.73

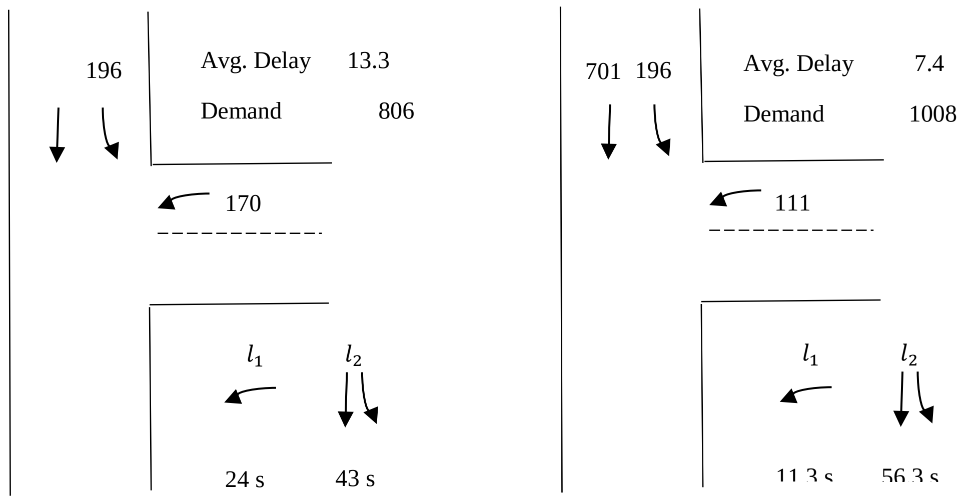
Table 4.18. Control Delay and Link Travel Cost

Demand	Network Total Control Delay (min)		Network Total Link Travel Cost (min)		Delta in Network Total Control Delay (%)	Delta in Network Total Link Travel Cost (%)
	Synchro	Prop. Model	Synchro	Prop. Model		
Peak	8021.38	6374.53	6486.22	6504.18	-20.53	0.28

In Table 4.13 and 4.15, the upper part of the tables show the hourly traffic volume approaching each lane group before implementing the proposed method in which empty cells indicate lane groups that do not exist. Similarly, the lower part of the tables provide the new traffic assignment result after implementing the proposed method.

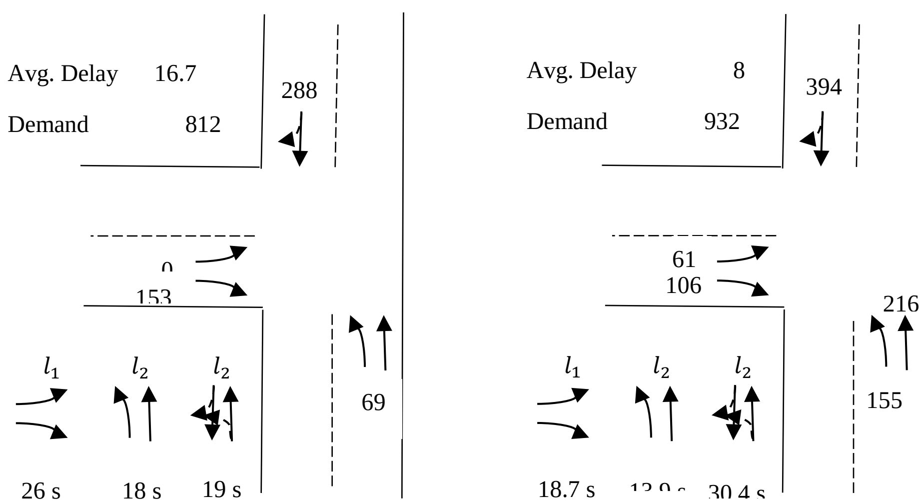
In peak-hour demand case, optimization results of the proposed model show significant improvements in both network total delay and total travel cost comparing with the optimization results from Synchro. A 20.53 percent of total delay reduction and an 11.23 percent of total travel cost reduction are observed in peak-hour demand case. Similar to what has been observed in the previous sample network, change in network total link travel cost before and after using the proposed model is not significant. As shown in Table 4.18, there is a 0.27 percent jump on total link travel cost after using the proposed model, and this increment in total link travel cost is caused by the traffic diversion after implementing the new signal settings in the system. Thus, reduction of total travel cost, which consists of control delay and link travel cost, is all contributed by intersection control delay reduction after the optimization. Comparing with the results in the previous sample network, demand in this case, even it is in AM peak, is more close to the low demand case which has around 30% V/C ratio at intersections. However, the improvement in this real world case is much more significant than that in the low demand sample network case. One major reason is the size and complexity of the network. Traffic diversion in the real world network is more common, as shown in Table 4.14, very few lane groups maintained same traffic flow before and after optimization. While in small network, as explained before, traffic entering the network sometimes had limited and incomparable path choices, what lead to little space for system improvement. Therefore, the proposed model would be more applicable and beneficial in larger network than small network with limited paths and intersections.

Moreover, the limitation of conventional signal timing optimization method, such as Synchro that has been using as example in this study, can be discovered in the result comparison as well. In Table 4.17, top three intersections that contribute the most in delay reduction are intersection 6, 9, and 12. Figure 4.4 to 4.6 illustrate detailed information of these three intersection before and after optimization. To more clearly show the changes and provide a more comprehensive look of the intersections, signal timings, phasing design, and traffic flows of all approaches of the same intersection are presented in one figure.



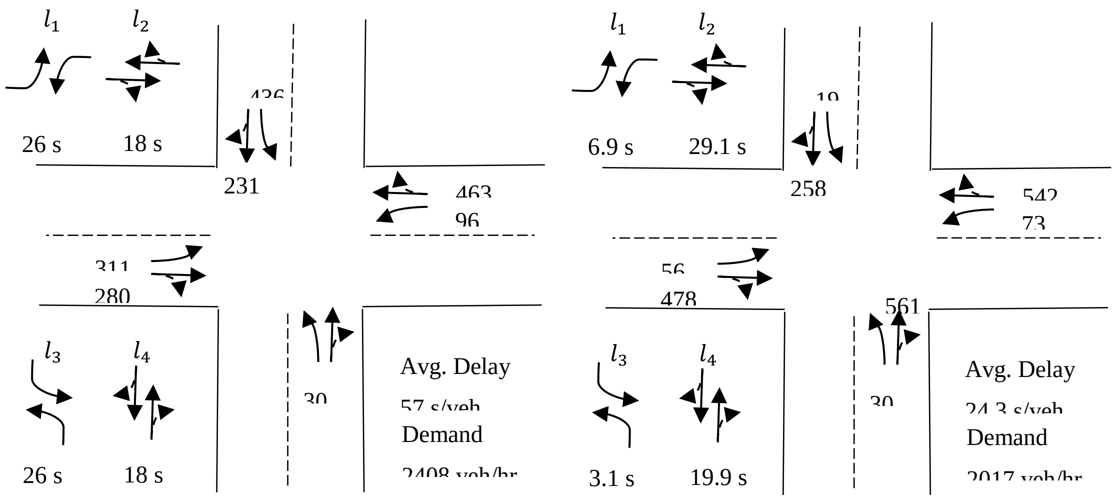
(a) (b)

Figure 4.4. Intersection 6 Results: Synchro (a) and MPEC (b)



(a) (b)

Figure 4.5. Intersection 9 Results: Synchro (a) and MPEC (b)



(a) (b)

Figure 4.6. Intersection 12 Results: Synchro (a) and MPEC (b)

As shown in Figure 4.4, at intersection 6, more green time was allocated to phase 2 after the optimization (43 seconds to 56.3 seconds), and, as a consequence of the change in system signal timings, some travelers changed their paths to the ones that traverse intersection 6 through southbound (440 veh/hr to 701 veh/hr) which has high capacity, and less travelers used the paths through westbound (170 veh/hr to 111 veh/hr). With all these changes, intersection average vehicle control delay decreased nearly 45 percent, from 13.3 s/veh to 7.4 s/veh, and even with increased traffic demands, 806 veh/hr to 1008 veh/hr, intersection total vehicle control delay still had over 30 reduction after optimization.

At intersection 9, as shown in Figure 4.5, the first two phases are split phasing, in which all lane groups of a particular approach move together. However, it can be easily observed that traffic assignment results before optimization was inappropriate because of the extreme imbalance of traffic load between left-turn lane group and through and right-turn lane group on eastbound and northbound. For instance, on eastbound, before optimization, no traffic was observed using left-turn lane group while 153 veh/hr was using through and right-turn lane group. However, after optimization, some travelers changed their routes and traffic loads on eastbound and northbound approaches of intersection 9 were balanced, which means that in the first two phases, intersection capacity could be utilized better. And similarly to intersection 6, even with increased traffic demands, 812 veh/hr to 932 veh/hr, intersection total vehicle control delay still had over 45 percent reduction after optimization in that vehicle average control delay was halved.

Situation at intersection 12 was different with what happened at intersection 9. Almost all of the improvement of intersection performance in terms of intersection control delay was contributed by the reduction of left-turn traffic. Left-turn traffic from all four approaches significantly decreased after optimization except the northbound which remained the same, and, in the contrast, through and right-turn traffic increased. These changes, combined with re-allocation of the signal timing plan which gave more green time to the through and right-turn phases, lead to reductions in both intersection average vehicle control delay and total traffic demands, and produced an over 64 percent reduction in total vehicle control delay at intersection 12.

From the observations from these intersections, it can be indicated that changes in the traffic routing was the main reason and power that caused the improvement in system performance, and is also the major difference between Synchro and the basic MPEC model proposed in this study.

In sparse demand case, an unexpected significant improvement, 36.81%, of total delay reduction can be observed in the result. Generally, in the case of extremely low demand, control delay is expected to be very low and does not play a vital role. There should not have big improvement after optimization in that the input traffic assignment had already been optimized based on link travel time only. This abnormally large improvement is caused by the bad performance of Synchro optimization in this case. Under sparse traffic demand condition, optimization results from Synchro have no improvements at most of the intersections. Although the significance of result comparison is lost because of this, it still shows the capability of the proposed model in dealing with sparse demand.

The computation time for both cases did not exceed 5 min CPU time, and it cost around 40 s CPU time for the sparse demand case and around 240 s CPU time for the peak demand case.

CHAPTER 5. THE ENHANCED MODEL

This chapter presents the problem statements, methodology development, and model formulation of the preliminary enhanced MINLP with Complementarity Constraints model. Genetic Algorithm and Enumerative Algorithm are employed in the attempt to solve the model.

5.1 Problem Statement

Signal phasing design refers to plan how vehicles that entering the intersection from all approach move in a signal cycle. Different vehicle movements may allow to move in the same phase only if conflicts do not exist between them. Then, simply speaking, a phase can be defined as a combination of movements that allow to move in a same time duration.

In practice, plan of traffic signal phases is highly depends on the design of signalized intersection and the traffic condition, including intersection's geometric design, channelization design, traffic signal control equipment, vehicle traffic demands, pedestrian traffic demands, and so on. This is reason that in this enhanced model, phasing designs are assumed to be pre-defined as candidate for selecting.

However, for normal intersections, signal phasing plans can be categorized into four general types by the sequence of protected left-turn movement and other movements: 1) lead-lead left-turn, 2) lag-lag left-turn, 3) lead-lag left-turn, and 4) splits phasing. Figure 5.1 illustrates examples for these four type of phasing designs.

Lead-lead left-turn is the most commonly used left-turn sequence which has both opposing left-turn movements start moving at the same time, and this protected, left-turn only phase always start before the phase for through movements on the same street. As the most commonly used design, the advantages of this phasing are: 1) drivers react quickly to the leading green arrow indication, 2) it minimizes conflicts between left-turn and through movements on the same approach, and 3) it cause less control delay in that put the protected phase before the permitted phase (normally the same phase with through movements on the same approach) makes the maximum queue length shorter.

Lag-lag left-turn phasing has both opposing left-turn end at the same time, while start after through movements on the same street. This type of phasing plan can offer operational benefits, but, as mentioned earlier, has some disadvantage when it is used with protected and permitted phase. Drivers who are waiting to make left-turn tend not to react quickly at the beginning of the phase. Moreover, since the maximum queue length is longer than other types, if a left-turn bay does not exist or is relatively short, then queued left-turn vehicles may block the inside through lane during the previous through movement phase.

Lead-lag left-turn phasing is generally used to accommodate through movement progression in a coordinated signal system. It benefits through movements on the major street which has dominant traffic volume.

At last, split phasing represents an assignment of the right-of-way to all movements of a particular approach, followed by all of the movements of the opposing approach. This type of phasing would be useful when the intersection has both high left turn and through volume or shared left turn and through lanes are used. In these situations, split phasing would more efficiently utilize the existing system.

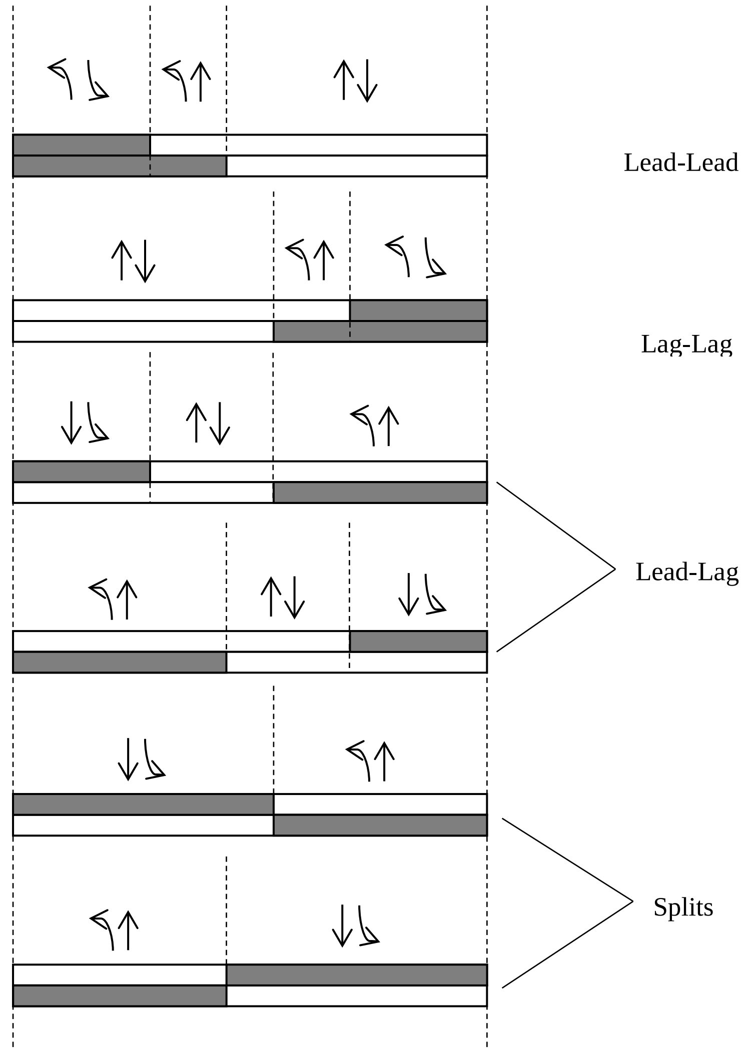


Figure 5.1. Typical Phase Settings

The basic MPEC model proposed and discussed in the previous two chapters successfully modeled the combined signal timing optimization and static traffic user equilibrium assignment problem by a rigorous MPEC model which is well applicable to real world problem, and can be efficiently and effectively solved by GAMS/NLPEC solver. In the enhanced model that is developed in this chapter, the author attempts to add selection of different phasing designs, which are predefined, into the basic model by employing binary variables to make the decision. Thus, the enhanced model is a mathematical problem that minimizes the travel cost, which consists of roadway segment travel time cost and signalized intersection control delay, in a given roadway transportation system by determining both the optimal effective green times and signal phasing design, while considers traffic user equilibrium assignment simultaneously.

5.2 Mathematical Model

Formulations and settings of roadway network, link, intersection, and signal settings in the enhanced model will continue to use those that have been defined in the basic model.

5.2.1 Phasing Design Selection. In addition, $\boldsymbol{N}_k$ is the set of nodes that are intersections with candidate phasing designs, and it is a subset of all signalized intersection in network, $\boldsymbol{N}_s$. Meanwhile, candidate phasing designs for an intersections are denoted as $\boldsymbol{K}_n$, which is the set of possible phasing designs for signalized intersection $n \in \boldsymbol{N}_k$, and its binary element k_n^m equals to 1 when phasing design m is selected for intersection n, 0 otherwise. $m = [1, 2 \dots]$ is defined to indicates different candidate phasing designs, and satisfies $m \leq M_n$, where M_n is the number of candidates that intersection n has. Since for one signalized intersection, only one phasing design can be selected at one time. k_n^m needs to satisfy the following condition.

$$\sum_m k_n^m = 1, \forall\, n \in \boldsymbol{N}_k$$

Then, using the binary variable k_n^m defined above, combination of signal phasing design for different signalized intersections could be presented by entirety of $\boldsymbol{K}_n$ for all $n \in \boldsymbol{N}_k$, and this combination is denoted as $\boldsymbol{K}$.

5.2.2 Model Formulation. The enhanced model is developed based on the basic MPEC model introduced in the previous two chapters. Similar to the basic model, this model also attempt to minimize the total vehicle travel time, including link travel time on roadway segments and intersection control delays at signalized intersections. Meanwhile, user equilibrium traffic assignment, time-dependent stochastic intersection control delay, mutual effect between these two, and necessary operational feasibility of signal settings are simultaneously considered in the model. In addition, the enhanced model further considered the potential benefits from different signal phasing designs, and use binary variable k_n^m to make selection for intersections. To avoid repetition, detail derivation process of the objective function, user equilibrium variational constraints, and delay calculation constraints will not be listed in this section, and it could be found in Chapter 3.

Complete model formulation pf the enhanced model is provided below followed by a brief explanation.

$$Min \sum_{p} f_p c_p$$

Subject to:

$$UE\left(f_p, c_p\right) = 0, \forall\, p \in \boldsymbol{P}, o \in \boldsymbol{O}$$

$$\sum_{l_n} \left(x_n^{l_n} + t_{lost}\right) = \omega_n, \forall\, n \in \boldsymbol{N}_s$$

$$c_p = \sum_{a} \delta_a^p h_a + \sum_{n \in \boldsymbol{N}_s, a \in \boldsymbol{A}_n, i} \gamma_{nai}^p d_{nai}\,, \forall\, p$$

$$d_{nai}(f_p, x_n^{l_n}, \boldsymbol{K}_n) = 0, \forall\, i, \forall\, a \in \boldsymbol{A}_n, \forall\, n \in \boldsymbol{N}_s$$

$$\sum_{m} k_n^m = 1, \forall\, n \in \boldsymbol{N}_k$$

$$gmin_n^{l_n} \leq x_n^{l_n} \leq gmax_n^{l_n}, \forall\, l, \forall\, n \in \boldsymbol{N'}$$

$$k_n^m = binary, \forall\, n \in \boldsymbol{N}_k, m$$

$$f_p \geq 0, \forall\, p \in \boldsymbol{P}$$

The objective function (25) minimizes the total travel cost in the roadway system. Equations (26) consist of UE variational inequality constraints that ensure UE assignment is always satisfied for any feasible solution of the model. Equation (27) ensures the feasibility of signal timing settings. Equation (28) gives the total travel cost per vehicle of one path as the summation of link travel time on every roadway segment and intersection control delay at every lane group that the path traverse. While, all nonlinear constraints and variational inequality constraints used in the calculation of intersection control delay are denoted as Equiation (29). Unlike the basic MPEC model, control delay is not calculated based on fixed phasing design but changes by different candidate phasing plans, $\boldsymbol{K}_n$. For one intersection, only one phasing design is chosen at one time and this is ensured by Equation (30). Moreover, Equations (31) bound the value of green splits in the pre-defined min-max range. Finally, Equation (32) ensures the values of path flow and binary variable k_n^m are feasible.

Introduction of binary variable makes the proposed model become a Mixed-Integer Nonlinear Programming (MINLP) with Complementarity Constraints, which is not able to be solve using GAMS/NLPEC solver directly. Two preliminary solution methods for the enhanced model will be tested.

5.3 Solution Methods

5.3.1 Genetic Algorithm. A GA method is proposed as one preliminary method to solve the enhanced model.

Chromosome, which defines a possible solution to the problem, would include a set of green times allocated to every phase for all signalized intersections, a set of traffic flows assigned to all possible paths, and a set of phasing design decision variable, k_n^m. Therefore, a possible solution of the model could be represented by the array shown below:

$$X = \left[x_1^1 \quad x_1^2 \quad x_1^3 \ldots x_2^1 \ldots x_n^{l_n} \vdots f_1 \quad f_2 \ldots f_p \vdots k_1^1 \quad k_1^2 \ldots k_2^1 \ldots k_n^m\right]$$

Most of the constraints and equations in the model can be transplanted into GA from the model formulation without any change. However, to represent UE constraints and control delay calculation, queue length estimation to be exact, variational inequality is not necessary nor applicable in GA representation. Variational inequality in UE constraints can be replaced by the equation shown as below.

$$f_p \cdot \left(c_p - \sum_o \theta_o^p \tau_o\right) = 0, \forall\, p \in \boldsymbol{P}$$

Meanwhile, vatiational inequality used for queue length estimation is not necessary anymore in that the maximum logic is able to be directly modeled in GA.

$$Q_{nai}^{l_n} = \max(0, Q_{nai}^{l_n - 1} - x_n^{l_n} \cdot w_{nai}^{l_n}), \forall\, l_n, \forall\, nai \in \boldsymbol{NAI}_1$$

Each component of the objective function, $f_p c_p$, could be treated as the product of traffic flow assigned on path f_p and the total travel cost of the path c_p. Thus, the objective of the model could be measured in GA as the fitness function FM.

$$FM = \sum_p f_p c_p$$

In this study, GA is set to terminate after 48 hour running time. While, the model normally does not start with a valid solution which satisfies all the constraints, and GA is supposed to be much more sufficient if the model start with a feasible solution. To avoid the unpredictable number of unnecessary trials at the beginning of the optimization, an initial valid solution is always obtained by special pre-optimization seeking the solution satisfying the maximum constraints in the model and terminated when all the constraints are met. Figure 5.2 shows the block diagram of proposed GA method.

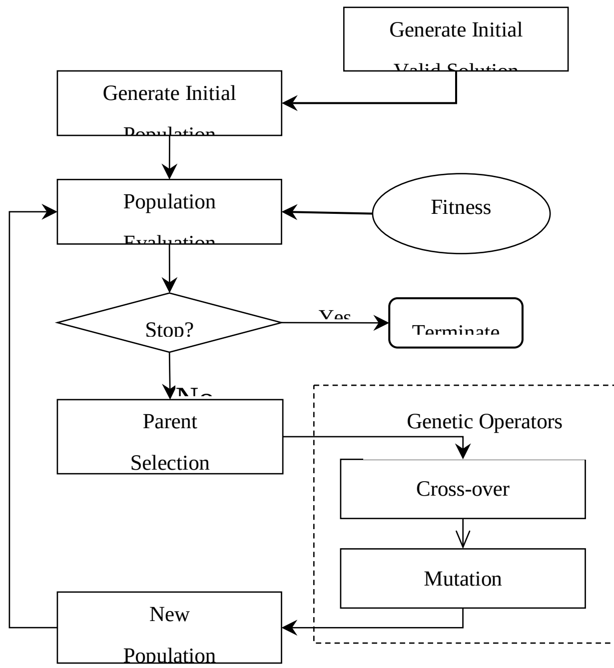


Figure 5.2. Genetic Algorithm Structure

5.3.2 Enumerative Algorithm. Another solution is also proposed and tested in the study. In the enhanced model, the reason of introducing binary variable to the basic MPEC model is to find the best phasing design for one or more signalized intersection from given candidates. And the model will provide different results for each combination of phasing designs selected for those intersections. But, obviously, the number of possible combinations is not countless. For instance, if the target roadway network has four signalized intersections, for which multiple phasing designs are given, and the number of candidate phasing designs, including the original signal settings, are 2, 3, 4, and 5, respectively. Then the total number of different combination in this network would be

$2 \cdot 3 \cdot 4 \cdot 5 = 120$. If only look at one combination out of all, it is nothing but the basic MPEC model, which can be solved by using NLPEC solver easily.

Therefore, if the possible combinations of candidate phasing designs are treated separately as single MPEC problems instead of one MINLP with Complementarity Constraints, Enumerative Algorithm (EA) can be employed to solve the enhanced model by solving a part or all of the single MPEC problems, and eventually, found the optimal solution of the enhanced model.

Both GA and EA method will be preliminarily applied in network 1 to test their capability of solving the enhanced model.

CHAPTER 6. COMPUTATIONAL EXPERIMENT USING THE ENHANCED MODEL

Both GA and EA solution method proposed in the Chapter 5 are tested to solve the enhanced model and implemented on a small network to verify the feasibility of the solution methods. Besides, another equally important purpose is to proof the assumption that changing phasing design is able to improve the performance of the entire signal controlled roadway system.

6.1 The Test Network

6.1.1 The Test Network and Settings. The small network that was used as network one in the computational experiment for basic MPEC model is going to be in this chapter.

It includes 4 signalized intersections, $[n_1, n_2, n_3, n_4]$, with pre-timed signal control and 8 origin/destination nodes around, $[o_1, o_2, \dots, o_8]$. Nodes in the network are connected by 24 directed links, $[a_1, a_2, \dots, a_{24}]$, each of which represents one direction of the roadway segment that connects intersections and/or origin/destination. Free flow travel time on any link in the network are assumed to be 15 seconds. All links that head to signalized intersections are modeled with two lane groups of traffic, which consists of one left-turn lane group i_1 and one through and right-turn lane group i_2. Figure 6.1 illustrates the small network described above.

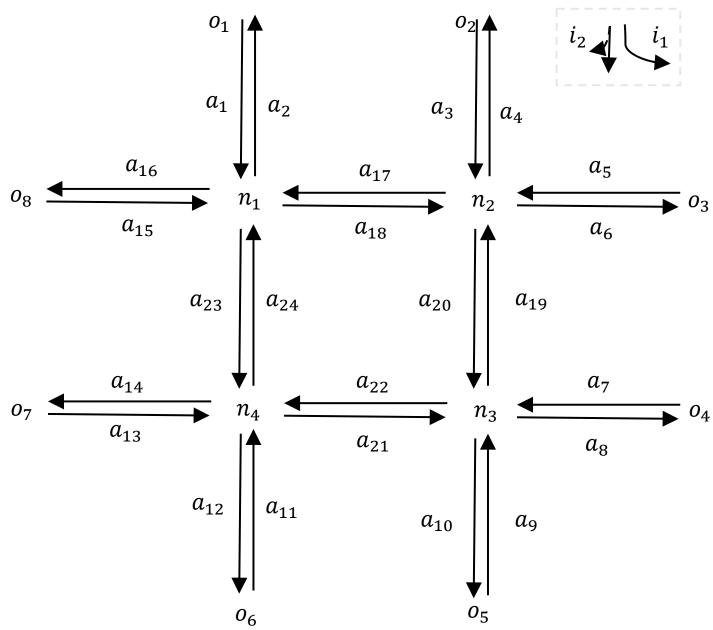


Figure 6.1. Experimental Network

As the original phasing design of the network, in order to test different types of intersection phasing, 3 different pre-timed plans were implemented. Since the use of lag-lag or lead-lag phasing design is mainly caused by operational benefits, both of them have similar or worse performance than lead-lead phasing when considering mobility (control delay) only. Thus, only lead-lead phasing and split phasing are implemented in the network. At intersection n_1, a default 4-phase protected-permitted lead-lead left-turn phasing is used, which includes a combination of protected left-turn phase, which only allows left-turns move, and permitted left-turn phase, which allows left-turn moves after yielding to conflicting traffic and pedestrians. While for intersection n_2 and n_4, protected-only left-turn phasing is employed as no permitted left-turn movement is allowed. For these three intersections, east west (EW) left-turns move in phase l_1, which is followed by moving of EW through movements in phase l_2. Similar process happens for north south (NS) approaches in phase l_3 and l_4. In contrast with the phasing plans used at other 3 intersections, a split phasing plan is employed at intersection n_3. In the split phasing plan, all movements from one approach are allowed to move simultaneously in one phase, in which no movement from other approaches is allowed to move. Figure 4 illustrates the intersection phase settings used in this small network case.

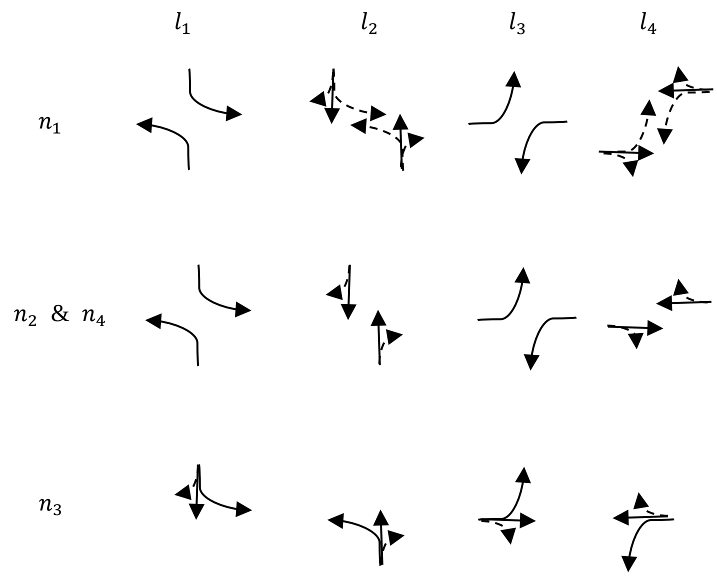


Figure 6.2. Original Phase Plans

Based on the results in Chapter 4, in Table 4.10, vehicles go through intersection 2 and 3 always experienced the longest intersection control delay in all three demand conditions. Especially for intersection 3, it can be observed that traffic entering the intersection from the westbound didn't change their route before and after optimization, and there existed an extreme imbalance between left-turn lane group and through and right-turn lane group. For instance, under high demand condition, left-turn traffic on westbound of intersection 3 was 162 veh/hr, while through and right-turn traffic was 1048 veh/hr. For this kind of traffic characteristic, split phasing may not be appropriate and optimal. Thus, in this computational experiment, two types of phasing designs are proposed for intersection 2 and 3, including a lead-lead left-turn phasing and a split phasing. Then, for two intersections, four different combinations, including the original phasing plan, are available in the optimization. Figure 6.3 illustrates all the possible phasing plans.

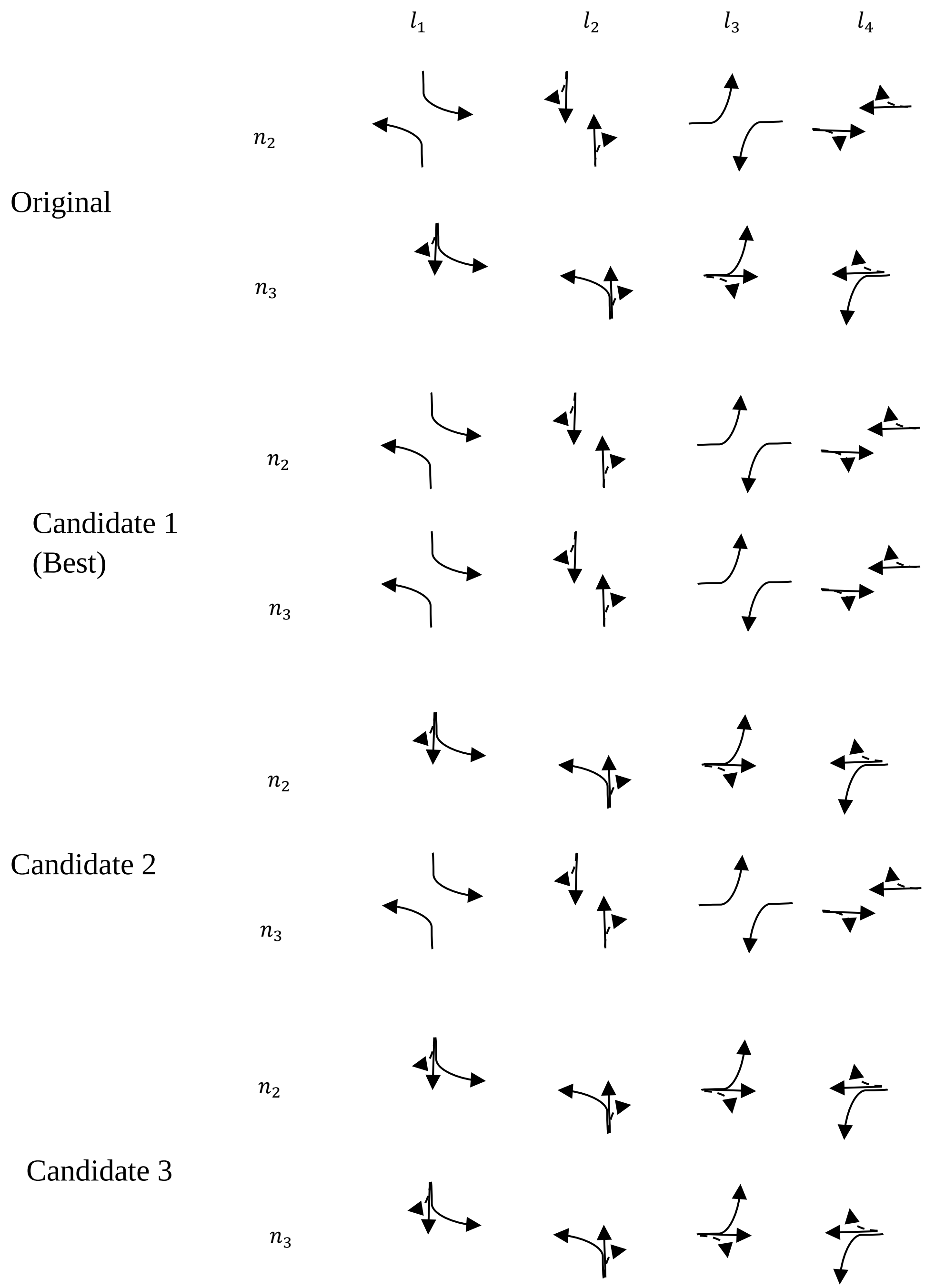


Figure 6.3. Candidate Phasing Design for Intersection 2 and 3

6.1.2 Test Method. Similar to application of the basic MPEC model, Three different traffic conditions, low demand, med demand, and high demand, are tested in the small network case with average V/C ratios ($\frac{v_{nai}}{\bar{s}_{nai}}$ in the model) for intersections range from 30% (low) to 150% (high). Traffic demands between OD pairs are randomly generated, and by this test method, performance of the proposed model under different traffic loads were tested. Detailed OD Demands for all three demand conditions could be found in Table 4.1, 4.2, and 4.3.

6.2 Results and Findings

6.2.1 GA method. In this study, GA of the proposed model was implemented and solved in an EXCEL-based solver, EVOLVER 6. It was successful implementing GA for the enhanced model, however the solving results were disappointing, or in another word, no valid optimization result was obtained.

Using the personal computer that also had been used for the basic MPEC model application, it took over 24 hours of running time to reached a feasible solution under the low demand condition by solving the pre-optimization model that maximize the number of satisfied constraints, while failed to gain any feasible solution for both medium or high demand cases.

Then, start from the feasible solution, it could not make any valid progress to another better feasible solution before it reached the termination condition, 48 hours of running time.

Failure in attempt to solve the enhanced model were caused by many reasons, one of which might be the limitation of computational power the personal computer has. It is undeniable that better results may be available after importing stronger computational resource, such as clusters. But, the tradeoff between solution efficiency and solution quality has to be considered. Comparing with the EA method, which uses the basic MPEC model to solve a part or all of the possible sub-problems with predefined phasing plans, surely the EA method requires more time in input preparation and pre-optimization process, however, it has a much better performance in computational time and is guaranteed a local optimal as well. With requirement of much more computational resource, which, as mentioned in previous chapters, may not be maintained in most of agencies, GA method provide a possibility of results with better or worse optimality, and it is not worthy most of the time.

Another reason is the on the model formulation itself. In this study, to reach traffic equilibrium, Static User Equilibrium (SUE) method was adopted and Wardrop's first principle of route choice was followed. Because of the SUE, in the bi-level structure of the proposed model, results of the upper level problem, which is signal setting of the system, and results of the lower level problem, which is traffic assignment of the system, are nearly one-to-one correspondence. For a fixed signal setting, number of traffic assignments that satisfies the SUE is limited. Hence, without good step direction, optimization of this bi-level problem is certainly inefficient. GA method may have a better performance if an equilibrium method that allows more tolerance is adopted. For the proposed model in this study, MPEC model has a better computational performance, and is able to provide acceptable results effectively and efficiently.

6.2.1 EA method. After solving the MPEC model for all four candidate phasing combinations as illustrated in 6.3, the best results on total travel time saving was observed in candidate 2, in which both of intersection 2 and 3 are implemented lead-lead phasing. Results of model application under three different traffic conditions are presented in the following tables. Table 6.1, 6.2, and 6.3 list the hourly traffic flow loaded on each lane group for Synchro optimization results, the basic MPEC optimization results with original phaisng, and the enhanced model results. Corresponding signal timings results are presented in Table 6.4 to 6.6. Meanwhile, detailed comparison of the optimization results is provided in Table 6.7 and Table 6.8.

Table 6.1. Low Demand Traffic Assignment Results (V/C = 30%)

f_{nai} (veh/hr)	EB		WB		NB		SB	
Intersection	LT	T&RT	LT	T&RT	LT	T&RT	LT	T&RT
(Synchro)								
1	36	148	68	209	112	266	115	151
2	169	194	58	144	101	295	61	194
3	76	256	54	356	122	133	101	94
4	212	248	90	234	11	108	65	220
(Basic)								
1	36	148	76	112	212	83	173	94
2	90	151	76	126	4	475	61	194
3	256	194	54	356	119	137	148	101
4	65	396	65	349	11	108	4	227
(New Phasing)								
1	36	148	16	112	212	281	162	104
2	173	256	110	92	4	277	61	194
3	57	205	54	356	119	137	137	160
4	227	233	140	333	11	108	14	167

Table 6.2. Medium Demand Traffic Assignment Results (V/C = 90%)

f_{nai} (veh/hr)	EB		WB		NB		SB	
Intersection	LT	T&RT	LT	T&RT	LT	T&RT	LT	T&RT
(Synchro)								
1	72	302	112	436	328	338	216	313
2	356	292	198	187	241	472	122	385
3	0	418	108	702	112	436	202	198
4	374	281	223	403	22	212	133	414
(Basic)								
1	72	302	208	292	388	236	258	271
2	239	269	163	222	98	700	122	385
3	229	329	108	702	248	299	208	185
4	155	500	142	532	22	212	88	510

(New Phasing)								
1	72	302	77	313	367	328	252	277
2	264	328	210	175	119	564	122	385
3	93	382	108	702	222	325	202	269
4	282	374	275	508	22	212	97	379

Table 6.3. High Demand Traffic Assignment Results (V/C = 150%)

f_{nai} (veh/hr)	EB		WB		NB		SB	
Intersection (Synchro)	LT	T&RT	LT	T&RT	LT	T&RT	LT	T&RT
1	108	461	212	590	421	565	342	464
2	464	454	169	410	302	785	184	569
3	76	626	162	1048	378	457	310	256
4	565	421	270	695	32	324	194	670
(Basic)								
1	108	461	132	483	528	362	481	325
2	378	490	249	330	195	988	184	569
3	279	473	162	1048	378	457	399	400
4	286	700	317	835	32	324	41	590
(New Phasing)								
1	108	461	79	540	472	478	360	446
2	429	429	310	269	252	872	184	569
3	163	608	162	1048	342	493	299	418
4	402	584	411	737	32	324	176	536

In Table 6.1 to 6.3, left turn traffic diversions from intersection 3 to intersection 4 can be obviously observed under all three demand conditions. Comparing the eastbound traffic volume at intersection 3 and 4 in the basic MPEC model results and the enhanced model results (highlighted with light blue color), it can be found that after selecting lead-lead left-turn phasing instead of the original split phasing at intersection 3, some left turn traffic switched their route to other paths, and as a consequence, volume on eastbound of intersection 4 was redistributed. Eastbound of intersection 4 is the main upstream of eastbound of intersection 3, since travelers tended to not make left turn on eastbound of intersection 3, some of them decided to make left turn earlier at intersection 4. This is why the left turn volume shifting happened. Meanwhile, for other approaches at intersection 3, not as significantly as eastbound, left turn volumes deceased while through and right turn volumes increased.

Matched with traffic diversion and redistribution, signal timing at intersection 3 also changed as shown in Table 6.4 to 6.6. More green time was given to phase 4 in which east- and west-bound through moves, while less green time was provided to phase 3 in which east- and west-bound left-turn moves. Then, this is more reasonable than the original phasing design and timing plan, and created significant improvement in the performance of entire system.

Table 6.4. Low Demand Signal Timing Results (V/C = 30%)

Intersection (Synchro)	Phase 1 (s)	Phase 2 (s)	Phase 3 (s)	Phase 4 (s)	Lost Time per Phase (s)	Cycle Length (s)
1	9	32	7	26	4	90
2	11	28	17	18		
3	11	14	21	28		
4	8	24	20	22		
(Basic)						
1	26.3	10.8	9.1	27.8	4	90
2	7.4	40.6	10.8	15.3		
3	13.5	12.5	22.4	25.6		
4	1.9	21.9	9.2	40.9		
(New Phasing)						
1	21.7	27.3	4.9	20.1	4	90
2	7.3	27.2	17.7	21.8		
3	13.3	19	8.1	33.6		
4	3.2	16.7	24.3	29.8		

Table 6.5. Medium Demand Signal Timing Results (V/C = 90%)

Intersection (Synchro)	Phase 1 (s)	Phase 2 (s)	Phase 3 (s)	Phase 4 (s)	Lost Time per Phase (s)	Cycle Length (s)
1	16	26	3	29	4	90
2	13	27	19	15		
3	8	21	17	28		
4	8	31	17	18		
(Basic)						
1	24.1	17.0	13.4	19.5	4	90
2	8.0	36.4	15.0	14.6		
3	10.9	14.7	15.4	33.0		
4	6.0	27.7	10.9	29.4		
(New Phasing)						
1	24.0	21.5	6.6	21.9	4	90
2	9	30.4	16.7	17.9		
3	13.8	17.1	6.6	36.3		
4	6.6	21.1	19.2	27.1		

Table 6.6. High Demand Signal Timing Results (V/C = 150%)

Intersection (Synchro)	Phase 1 (s)	Phase 2 (s)	Phase 3 (s)	Phase 4 (s)	Lost Time per Phase (s)	Cycle Length (s)
1	14	29	5	26	4	90
2	11	30	17	16		
3	10	16	18	30		
4	7	24	20	23		
(Basic)						
1	27.3	16.8	6.6	23.2	4	90
2	9.2	33.2	14.5	17.2		
3	15.2	16.1	14.7	27.9		
4	1.9	21.8	15.3	35.0		
(New Phasing)						
1	22.0	22.5	5.2	24.3	4	90
2	10.6	31.2	17.3	14.9		
3	14.7	18.5	7.5	33.3		
4	6.6	19.5	19	28.9		

Table 6.7. Results Comparison

Demand	Int	Avg. Delay per vehicle (s)		Delta in Avg. Delay per vehicle (%)	Delta in Intersection Total Delay (%)	Delta in Network Total Delay (%)	Delta in Network Total Cost (%)
		Basic	New Phasing				
Low	1	31.3	29.1	-7.03	6.61	-5.08	-3.58
	2	35.3	38.9	10.20	9.26		
	3	47.9	30.2	-36.95	-43.42		
	4	26.8	38	41.79	42.72		
Medium	1	52.4	44.5	-15.08	-16.71	-20.37	-17.46
	2	81	70.8	-12.59	-13.83		
	3	131.6	88.5	-32.75	-32.90		
	4	70.2	66.7	-4.99	-5.51		
High	1	104.2	115.6	10.94	13.41	-12.54	-11.77
	2	240.6	221.6	-7.90	-9.78		
	3	366.1	247.4	-32.42	-33.61		
	4	179.4	208.4	16.16	19.03		

Table 6.8. Control Delay and Link Travel Cost

Demand	Network Total Control Delay (min)		Network Total Link Travel Cost (min)		Delta in Network Total Control Delay (%)	Delta in Network Total Link Travel Cost (%)
	Basic	New Phasing	Basic	New Phasing		
Low	2816.60	2673.52	1174.52	1174.52	-5.08	0.00
Medium	12328.13	9817.39	2173.81	2152.21	-20.37	-0.99
High	49852.77	43601.16	3248.71	3248.62	-12.54	0.00

In all three cases, comparing with the optimization results from the basic MPEC model, the proposed enhanced model did improve system performance as reducing both total control delay in the network and total travel cost which consist of control delay and link travel cost in the network. While the improvements differ under various traffic conditions. When demand is low, system performances after optimization using basic MPEC model and enhanced model ware not different significantly as improvement of 5.08% in delay reduction and 3.58% in total travel cost reduction.

The proposed model provides the best improvement in the case with medium demand (20.37% in delay reduction and 17.46% in total travel cost reduction) and a decent improvement with high demand (12.54% in delay reduction and 11.77% in total travel cost reduction). Unlike the control delay, total link travel time almost remained the same and might increase or decrease. As shown in Table 6.8, changes in network total link travel cost before and after using the proposed model did not exceed 1 percent.

All findings mentioned above strongly support our assumption that adding phasing design as a variable in the model will further create potential improvement of the system performance.

CHAPTER 7. SUMMARY AND CONCLUSION

7.1 Summary and Concluding Remarks

In this study, intersection control delay calculation method introduced in HCM 2010 has been employed in a combined optimization problem for area traffic signal control and network traffic assignment, and formulated as Variational Inequality (VI) constraints in the basic MPEC model. It allows the proposed method to accurately model and estimate the intersection control delay of various type of movements in real world scenarios such as those with multiple green phases and multiple control methods (protected, permitted, or mixed) without the use of simulation-based traffic model. The combined problem was formulated as mathematical programming with equilibrium constraints (MPEC) and solved by using GAMS/NLPEC solver which reformulates and solves the MPEC problem as standard nonlinear programming (NLP).

The basic MPEC model was applied on an experimental 4-intersection network and a real world problem with 13 signalized intersection in the City of Chicago urban area. Different phasing plans were adopted in the experimental network, and three traffic loads were tested as different cases from low traffic demand condition case (with intersection V/C around 30%) to high traffic demand condition case (with intersection V/C around 150%). Comparing the optimization results of the proposed model with the optimization results by using Synchro with the same initial traffic assignment, improvements in both total intersection control delay and total travel cost were observed in all three cases, and they varied significantly. Small improvement, 2.55% in total travel time reduction, was obtained in the low demand case, and large improvement, 14.54% in total travel time reduction, was showed in the medium demand case which has near capacity traffic loads at signalized intersections. After the optimization, drivers tended to switch their route from intersections with protected only left-turn phasing to intersections with protected-permitted left-turn phasing and split phasing, where more left turn traffic would better utilize the intersection capacity. Comparing with the protected left-turn only phasing, protected-permitted left-turn phasing and split phasing had relatively more capacity without occupying the green time for other phases.

For the real world problem, named as network two, two different OD demands generated by Chicago TRANSIMS microscopic traffic simulation model were tested. In the case with AM peak traffic demand, which is roughly 30% V/C, 11.23% total travel time reduction was obtained from the proposed method when compared with Synchro optimization result, and almost all of the travel time reduction was contributed by reduction in intersection control, 20.53%, in that network total link travel time remained basically the same with 0.28% increase. Under similar demand condition, 30% V/C ratio, the basic MPEC model tend to be more applicable and beneficial in larger network than small network with limited paths and intersections. Besides, it was also observed that changes in the traffic routing was the main reason and power that caused the improvement in system performance, and is also the major difference between Synchro and the basic MPEC model proposed in this study. However, in the case with off peak traffic demand, although the significance of result comparison was lost because of the bad optimization results from Synchro, it was still able to present the capability of the proposed model when dealing with extremely low demand situation.

Furthermore, in order to import more reality to the basic model and consider the potential system benefit that comes from different signal phasing designs, an enhanced model is developed based on the basic MPEC model by employing binary variables to make selection of optimal signal phasing plans from pre-defined candidates. The enhanced model belongs to a new class of challenging optimization problems, namely Mixed-Integer Nonlinear Programming (MINLP) with Complementarity Constraints. Formulating the problem with binary variables allows for the selection of proper phasing design, however, also increase the difficulty to solve the problem. As preliminary solution attempts, two heuristic solution algorithms, GA method and EA method, are proposed.

Both GA and EA solution method were implemented on the test network to verify the feasibility of the solution methods. In the network, one lead-lead left turn phasing and one split phasing designs were prepared as candidates for intersection 2 and 3, respectively. In total, 4 different combination of phasing plans were available in the problem.

Among two preliminary solution methods, GA failed to provide valid nor optimal solution within valid running period. While EA method, which highly relies on the basic MPEC model, provided optimal results when keeping original phasing at intersection 2 unchanged and replacing the split phasing at intersection 3 with normal lead-lead phasing. Comparing with the optimization results of the original phasing plans, 3.58%, 17.46%, and 11.77% reduction in network total cost were observed under low, medium, and high traffic demand conditions, respectively. Similar to previous cases, all reduction came from the improvement at signalized intersections, particularly, from intersection 3. The results strongly supported our assumption that adding phasing design as a variable in the model would further generate potential improvement in the system.

7.2 Future Research Direction

The application of the basic MPEC model, along with the solution method, does not require extensive data collection, preparation, and computational efforts as compared to the methods that rely on simulation-based traffic models to evaluate the performance of traffic signals. This gives it potentially greater applications to agencies that do not maintain rich data on travel demand, facility preservation, traffic operations, data processing and preparation capacity, and high performance computing facilities. However, current solution method relies on a good initial point to obtain an acceptable optimization result, and it would be useful to develop a better method to find a good initial point or initial feasible solution as future work.

For the enhanced model, an efficient solution algorithm is still under development. Both of the proposed preliminary solution methods have their limitations and required more research. Looking for an alternative of SUE, which allows more tolerance when locating feasible solutions, could be a future research direction for the GA method approach. Meanwhile, for EA method approach, a reduction method, which is able to effectively reduce the size of candidate phasing design combinations without losing solution optimality, are also needed to improve method’s efficiency.

BIBLIOGRAPHY

Abdelfatah, A. S. & Mahmassani, H. S. (1998). System Optimal Time-Dependent Path Assignment and Signal Timing in Traffic Network. *TRR* 1645, pp. 185–93.

Abdelfatah, A. S. & Mahmassani, H. S. (2001). A Simulation-Based Signal Optimization Algorithm within a Dynamic Traffic Assignment Framework. *in Proceedings of 2001 IEEE Intelligent Transportation Systems Conference*, Oakland, CA, pp. 428–33.

Allsop, R. E. & Charlesworth, J. A. (1977) .Traffic in A Signal Controlled Road Network: An Example of Different Signal Timings Inducing Different Routeings. *Traffic Engineering and Control*, 18 (5), pp. 262–264.

Allsop, R. E. (1974). Some Possibilities for Using Traffic Control to Influence Trip Destinations and Route Choice. *In D. J. Buckley (Ed.), Proceedings of the Sixth International Symposium on Transportation and Traffic Theory*, Sydney, Australia, 345-374, Elsevier, Amsterdam.

Cantarella, G. E., Improto G., & Sforza A. (1991). An Iterative Procedure for Equilibrium Network Traffic Signal Setting. *Transportation Research Part A*, 24, 241–249.

Ceylan, H. (2006). Developing Combined Genetic Algorithm Hill-Climbing Optimization Method for Area Traffic Control. *Journal of Transportation Engineering* 132 (8), 663–671.

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

Ceylan, H. & Bell, M. G. H. (2005). Genetic Algorithm Solution for the Stochastic Equilibrium Transportation Networks under Congestion. *Transportation Research Part B* 39 (2), 169–185.

Ceylan, H. & Ceylan, H. (2012). A Hybrid Harmony Search and TRANSYT hill climbing algorithm for signalized stochastic equilibrium transportation networks. *Transportation Research Part C* 25, 152–167.

Chiou, S. W. (1999). Optimization of Area Traffic Control for Equilibrium Network Flows. *Transportation Science*, 33, 279-289.

Cohen, S. L. & Mekemson J. R. (1985). Optimization of Left-Turn Phase Sequence on Signalized Arterials. *Transportation Research Record*, Vol. 1021, 53-58.

Dickson, J. J. (1981). A Note on Traffic Assignment and Signal Timings in a Signal-Controlled Road Network. *Transportation Research*, 15 B, 267-272.

Dion, F., Rakha, H., & Kang, Y.S. (2004). Comparison of Delay Estimates at Under-Saturated and Over-Saturated Pre-Timed Signalized Intersections. *Transportation Research part B*, Vol. 38, 99-122.

Friesz T. L., Tobin R. L., Cho H. J. & Mehta N. J. (1990). Sensitivity Analysis Based Heuristic Algorithms for Mathematical Programs with Variational Inequality Constraints. *Mathematical Programing*, 48, 265-284.

Gartner, N. H. & Al-Malik, M. (1996). Combined Model for Signal Control and Choice in Urban Networks. *TRR* 1554, pp. 27–35.

Gartner N. H., Gershwin S. B., Little J. D. C., & Ross P. (1980). Pilot Study of Computer-Based Urban Traffic Management. *Transportation Research*, 14B, 203-217.

Heydecker, B. G. & Khoo, T. K. (1990). The Equilibrium Network Design Problem. *in Proceedings of AIRO'90 Conference on Models and Methods for Decision Support*, Sorrento, Florida, pp. 587–602.

Homburger, W. S. (1982). *Transportation and Traffic Engineering Handbook, Second Edition*. Prentice-Hall.

Hunt, P. B., Robertson, D. I, Bretherton, R. D. & Winton, R. I. (1981). *SCOOT - a traffic responsive method of co-ordinating signals*. TRRL Laboratory Report 1014, Road Research Laboratory, England.

Meneguzzer, C. (1995). An Equilibrium Route Choice Model with Explicit Treatment of the Effect of Intersections. *Transportation Research Part B*, 29 (5), pp. 329–356.

Robertson, D. I. (1969). *TRANSYT: A traffic network study tool*. RRL Report LR

253, Road Research Laboratory, England.

Roess, R. P., Prassas, E. S., & Mcshane, W. R. (2004). *Traffic Engineering, Third Edition*, Perason Education, Inc.

Sheffi, Y. & Powell, W. B. (1983). Optimal Signal Setting Over Transportation Network. *Journal of Transportation Engineering*, 109, 824–39.

Smith, M. J. (1980). A Local Traffic Control Policy Which Maximizes the Overall Travel Capacity of an Urban Road Network. *Traffic Engineering and Control*, 21, 298–302.

Smith, M. J. (1981a). The Existence of an Equilibrium Solution to the Traffic Assignment Problem When There Are Junction Intersections. *Transportation Research*, 15B, 443–51.

Smith, M. J. (1981b), Properties of a Traffic Control Policy Which Ensure the Existence of Traffic Equilibrium Consistent With the Policy. *Transportation Research*, 15B, 453–62.

Strong, D. W. & Rouphail, N. M. (2006). Incorporating the Effects of Traffic Signal Progression into the Proposed Incremental Queue Accumulation (IQA) Method. *Proceedings of the 85th annual meeting of the Transportation Research Board*, Washington, D.C.

Tian, Z., Mangal, V., & Liu, H. (2008). Effectiveness of Lead-Lag Phasing on

Progression Bandwidth. *Transportation Research Board*, Vol. 2080, 22-27.

Tobin R. L. & Friesz T. L. (1988). Sensitivity Analysis for Equilibrium Network Flows. *Transportation Science*, 22, 242-250.

Webster, F. V. (1958). Traffic Signal Settings. *Road Research Technical Paper 39*, Road Research Laboratory, London.

Yang, H. & Yagar, S. (1995). Traffic Assignment and Signal Control in Saturated Road Network. *Transportation Research*, 29A (2), 125–139.

Yin, Y. (2000). Genetic-Algorithms-Based Approach for Bilevel Programming Models. *Journal of Transportation Engineering*, 126(2), 115–20.