This is to certify that the
dissertation entitled

DEVELOPMENT OF DYNAMIC REAL-TIME
INTEGRATION OF TRANSIT SIGNAL PRIORITY IN
COORDINATED TRAFFIC SIGNAL CONTROL SYSTEM
USING GENETIC ALGORITHMS AND ARTIFICIAL
NEURAL NETWORKS

presented by

MOHAMMAD SHAREEF GHANIM

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in Civil Engineering

08/01/2008

Major Professor's Signature

June 6, 2008

Date

DEVELOPMENT OF DYNAMIC REAL-TIME INTEGRATION OF TRANSIT SIGNAL PRIORITY IN COORDINATED TRAFFIC SIGNAL CONTROL SYSTEM USING GENETIC ALGORITHMS AND ARTIFICIAL NEURAL NETWORKS

By

Mohammad Shareef Ghanim

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Civil Engineering

2008

ABSTRACT

DEVELOPMENT OF DYNAMIC REAL-TIME INTEGRATION OF TRANSIT SIGNAL PRIORITY IN COORDINATED TRAFFIC SIGNAL CONTROL SYSTEM USING GENETIC ALGORITHMS AND ARTIFICIAL NEURAL NETWORKS

BY

Mohammad Shareef Ghanim

Many transit agencies are interested in employing transit signal priority (TSP) at signalized intersections in urban areas to reduce travel time and achieve other operational benefits. However, state-of-the-art systems do not consider the influence of bus dwelling activities at bus stops, especially when the stops are located near the intersection stop line. They are also incapable of efficiently considering simultaneous priority requests and implementing TSP within coordinated traffic networks. This research develops a dynamic real-time control logic considering the influence of transit dwelling activities on priority requests and surrounding traffic performance. The problem of accurately predicting transit travel time in situations in which dwelling activities may take place before buses reach the stop line is first solved through the development of an Artificial Neural Network prediction model. The optimization of the prioritized and coordinated network traffic signals is then performed by using Genetic Algorithms. The resulting signal controller is tested in a simulated traffic network using the VISSIM microscopic traffic simulation software. The simulation results show that the proposed signal control model has the ability to improve transit and non-transit traffic operational performance. Test results more specifically show an ability to reduce delays and the number of stops incurred by transit vehicles, while improving schedule adherence and minimizing negative impacts on traffic resulting from the provision of transit priority treatments.

DEDICATION

I dedicate this dissertation to my parents, Shareef and Amal, my first teachers. Without their caring support and encouragement, I would have never been in the graduate school pursuing the graduate studies.

I dedicate this work to my beloved wife Amal. The care, patience, and encouragement she offered have a great influence in completing this dissertation.

To the person who does not know what is going on at this moment, my son Ahamd, I dedicate this work. I would like to thank him for his understanding. Even though he was a little baby when this research was completed, he knew when to quit disturbing me.

I also dedicate this work to my brother Ahmad, and my sisters, Heba, Waf'a and Reem, who are always happy for me and my success.

ACKNOWLEDGMENT

During the work on this research, there is no doubt about who gets my sincere thanks. I would like to thank my advisors, Dr. Francois Dion and Dr. Ghassan Abu-Lebdeh. Needless to say that without their guidance and directions, this work would not have been finished.

I would like also to thank Dr. William C. Taylor and Dr. Erik Goodman for joining the Ph.D committee, and for the valuable advices and discussions we had throughout my staying in MSU.

Obviously I would like to extend my thanks to Dr. Ron Harichandran, Mary Mroz, Laura Taylor, and Margaret Conner for all the help they offered to go through the complications I faced.

I am also grateful for the faculty and staff who helped me in my undergrad studies at An-Najah National University for their support and encouragement, Dr. Sameer abu-Eisheh, Dr. Khalid Al-Sahili, Dr. Najeh Tamim, Dr. Abdelrazaq Tuqan, Dr. Ryad Abdelkareem, Dr. Anan Jaioussi, Dr. Shaker Al-Bitar, Dr. Mohammad Ghazzal, Dr. Hafiz Shaheen, Dr. Muneer Abdo, Dr. Nehaya Awartani, Hamees Etbeeleh, Haitham Awadeh., Samia Al-Shanti, and Ahmad Al-Ahmad.

For many reasons that I cannot count, I am thankful to my friend Monther Dwaikat. Although a few words are not enough to do so, but this is the available space I have.

My sincere and deep thanks go to those whom we spent long time since childhood, Basel Ghanim, Nidal Abdullah, Akram Yameen, Murad Faris, and Ahmad Tawfeeq.

I am thankful for Hui Chen and Abrar Siddiqi for making my stay enjoyable and active at MSU. I appreciate the time we spent discussing and talking about several things outside the academic college life.

There are many people I have met during my academic and social life, and the conversations we had affect me in one way or another, therefore, I would like to extend my sincere thanks to them, Mohammad Halloush, Osama Ababneh, Hussein Kurt, Suhaib Salawdeh, Nassar Odwan, Ma'en Srouji, Tareq Zeyad, Tareq Ayyad, Kais Salhout, Mahmoud Dwaikat, Buran Khalid, Mahmood Khassawneh, Qasem Al-Zoubi, Salama Fayyad, Fuad Abu-Jarad, Hassan Abdelkareem, Ahmad Masa'deh, Nazmi Arar, Jomana Hamadenah, Ala'a Olian, Maya Abu-Elhayyat, Ibraheem Nasrallah, Rami Shar'ab, Haya Al-Aghbar, Jawad Al-Adham, Fawz Kobary, Maha Al-ruzzi, Samah Sarrawi, Ghorood Awadeh, and Suha Abu-Ebaya.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I.

INTRODUCTION

In response to increasing traffic congestion, many transit agencies are considering providing buses with preferential treatments along urban arterials. A frequent solution is to provide priority of passage to transit vehicles at signalized intersections through temporary signal timing adjustments designed to reduce incurred delays. Since intersection delays typically account for 15% to 20% of bus travel times [1], such adjustments offer, in theory, a significant potential for improving transit service along urban arterials through reduced commute times.

The primary objective of Transit Signal Priority (TSP) systems, as preferential systems are commonly referred to, is to reduce transit travel times by shortening or eliminating the time spent by vehicles waiting for a green signal to appear or a queue to dissipate. Achieved delay reductions can further lead to improved schedule adherence and help better regulate inter-arrival times, and thereby, contribute to increases in reliability and quality of service. By favoring the movement of high-occupancy vehicles over low-occupancy vehicles, TSP also has the potential for improving overall intersection efficiency on a per person basis. Local transit agencies can further benefit from induced ridership if motorists decide to switch to transit as a result of the improved service. In turn, such a modal shift may improve traffic operations at signalized intersections by reducing the overall traffic demand. TSP may finally provide cost savings to transit system operators by allowing a specific service to be offered with less person-hours of work or through the use of fewer vehicles, not counting the possibility of reducing the frequency of maintenance work on vehicles through reduced usage.

Although traffic signal enhancement strategies favoring the movement of transit vehicles through signalized intersections have been considered and applied since the 1970s [2], only recently has the use of such strategies increased. This is strongly linked to the availability of new technologies allowing improved vehicle detection, reliable communication between vehicles and roadside equipment, and enhanced traffic signal control functionalities. The majority of TSP applications are currently primarily developed in areas with extensive transit service (areas typically characterized by short headways and/or heavy ridership) and where moderate-to-heavy traffic conditions combined with the presence of traffic signals create significant impediments to transit service. However, because of its potential operational benefits and reduced cost of required technologies, TSP deployments are increasingly being considered for locations where traffic impediments may not be as significant as past applications [3, 4].

TSP is different from what is commonly referred to as signal preemption. While both processes result in temporary signal alterations, a fundamental difference exists in how these alterations are granted. Signal preemption is usually offered only to emergency vehicles, such as ambulances and fire trucks, and trains approaching railroad crossings. In this case, the signal alterations are generally implemented without consideration of the potential impacts on other traffic. The primary objective here is simply to switch the signal to green as soon as possible on the approach on which the emergency vehicle or train is located to reduce the risks of collision with conflicting traffic. On the other hand, signal priority is primarily provided to facilitate the movement of prioritized vehicles in mixed-traffic conditions. In this case, signal timing alterations are only provided within a certain set of conditions and when their impacts on other traffic are within tolerable limits. In most systems, signal preemption calls from emergency vehicles would override priority calls from transit vehicles.

While TSP has been shown to provide operational improvements to transit vehicles, it may negatively impact other vehicles in the traffic stream or pedestrian safety. One of the most commonly cited negative impacts is the possibility for vehicles on competing approaches to incur significant delay increases as their green time is being reduced to serve the approaching transit vehicle. This creates a need for TSP strategies to take into consideration the challenge of balancing the specific needs of transit vehicles and other users of the road.

This research focuses on using a genetic algorithm (GA) optimization method and artificial neural network (ANN) model to develop a real-time TSP controller for use in signal controllers operating within coordinated signalized urban networks. When compared with existing practice, the objective of the research is to develop a TSP logic providing enhanced operational performance when compared to current practice.

The remainder of this dissertation is divided into nine chapters organized as follows:

- The next section, Chapter 2, provides general information about TSP systems, including typical components and commonly used signal alteration strategies.

- Chapter 3 focuses on the problems associated with the implementation of TSP within signalized networks.

- Chapter 4 presents the objectives of this research and the general approach that has been adopted for solving the problem defined in Chapter 3. It also details the optimization and prediction methods that will be followed, as well as the simulation test-bed that will be used for developing and evaluating the proposed TSP strategies.

- Chapter 5 presents the results of simulation analyses that were conducted to assess the need for the development of an improved bus arrival time prediction tool considering situations in which bus stops located downstream of bus detection points can create difficulties in accurately predicting intersection arrival times .

- Chapter 6 presents the control variables that are considered in the provision of TSP and details the mathematical derivation of the general cost function that will be used by the genetic algorithms to determine near-optimal signal timings to effectively serve both transit and traffic needs.

- Chapter 7 presents a stop line arrival time prediction tool that has been developed based on the results of Chapter 5 for providing reasonably reliable information about the expected arrival time of buses at an intersection stop line on approaches on which dwelling activities downstream of a detection point may create significant uncertainties in predicting stop line arrivals.

- Chapter 8 details the experimental setup for evaluating the proposed traffic signal control algorithm against pre-timed signal control with and without active TSP. This Chapter also discusses the architecture of the optimization model that has been developed.

- Chapter 9 presents different measures of effectiveness that have been selected to assess the benefits associated with implementing the proposed traffic signal control algorithm against more traditional control strategies under various traffic conditions and transit demands.

- Chapter 10 highlights the main findings in this research and the main recommendations for extending the work in future research.

CHAPTER II.

PRIMER ON TRANSIT SIGNAL PRIORITY SYSTEMS

As indicated in the introduction, TSP systems attempt to provide some priority service opportunities to transit vehicles by implementing temporary signal timing alterations designed to reduce vehicle wait time at signalized intersections. For readers less familiar with such systems, this section presents an overview of traffic signal control systems and current practice associated with TSP systems. The following sub-sections detail more specifically the general objectives behind the provision of signal priority, typical system components, and frequently used strategies for signal timing alterations.

2.1. Traffic Signal Control System

Traffic signal control is one of the most effective methods for managing traffic at intersections. Signal control systems try to manage conflicts between various traffic streams by allocating passage rights in a specific sequence. The sequence by which vehicles are allowed to enter an intersection can be set to satisfy various objectives. The most common optimization objectives are to minimize vehicle delay, minimize the number of stops incurred, or maximizing vehicle throughput.

While traditional traffic signal control systems rely on historical data to develop optimal signal timing plans, recent systems increasingly rely on real-time data to provide adaptive control that automatically adjust to changes in traffic conditions. Regardless of the type of system used, traffic signal control might be approached from three different levels: isolated intersection level, arterial level, and network level. Controlling at the isolated intersection level occurs when the traffic signals simply aims to operate the traffic movements at one intersection, without any communicating with surrounding intersections. Control at the arterial level occurs when the intersections along a corridor are synchronized to operate together, either through the establishment of local communication links or through the use of a centralized control unit. Control at the network level finally occurs when intersections along different corridors are synchronized to optimize traffic flow throughout an entire sector or region.

There are some terminologies that are commonly used in designing traffic signal control systems that should be defined before presenting the different types of signal controls. These terms are cycle length, phase, phase split, and offset:

- Cycle length is the total time in seconds required for a complete sequence of signal indication color.
- Phase is the part of the cycle that is allocated for a specific traffic stream movement, or a combined traffic streams.

- Phase split is the ratio between the green time assigned for a given phase to the non-green time within a particular cycle.
- Finally, the offset is the time elapsed in seconds (or percentage) between the beginning of a given phase in one signalized intersection, and the beginning of the corresponding phase in the downstream signalized intersection.

Traffic signal control systems can vary in their degree of complexity and required technology. This has important implication on how TSP can be provided at an intersection. Existing systems can be classified under one of the following types:

Pre-timed Traffic Signal Control

In pre-timed signal control, the signal plan, such as cycle length, splits, and phases are based on historical traffic data and do not react to the traffic fluctuation. Pre-timed signal control is appropriate for areas where traffic demand is very predictable.

Actuated Signal Control

Actuated signal control systems allow the signal controller to provide some response to fluctuations in vehicle demand and to actuations by pedestrians. This type of system allows the controller to adjust the phase lengths or skip phases from a pre-timed signal plans to better match prevailing demands. In this case, the green time is a function of the number of service calls placed by approaching vehicles and pedestrians. The resulting phase length could be anywhere between pre-defined minimum and maximum green times. In all cases, the changes in phase duration are only temporary, as the system always reverts back to a predefined baseline fixed-time plan at the end of each cycle. As such, this approach thus only allows for adjustments around a predefined average demand.

Actuated signal control systems can be categorized into two main groups: semi-actuated and fully-actuated systems. In semi-actuated systems, all phases can be subject to actuation except the main green phase, the one serving the approach with the highest flow. These systems are commonly used along coordinated arterials, where it is desired to maintain a certain pattern of green indications between intersections to preserve specific flow progression patterns. In fully-actuated systems, all phases can have their duration adjusted to traffic demands. These systems are used at isolated intersections with balanced random demand from all approaches.

Traffic Responsive Signal Control

Traffic responsive signal control system uses information about current traffic conditions at an intersection or a group of intersections to select appropriate pre-define timing plans from a library of predefined plans. As such, these systems do not dynamically optimize signal timings. They simply use available information to determine when might be an optimal time to switch from one plan to another.

Adaptive (Real-Time) Signal Control

Adaptive (real-time) signal control systems are the most advanced types in operation today. These systems heavily rely on traffic detection to provide signal timing plans that vary with observed changes traffic demands. These systems collect information about approaching traffic and use this information to develop short-term signal timing plans that optimized for the expected flows. New signal timing plans can be developed by either modifying an existing plan or generating an entire new plan. Re-optimization can occur once every few minutes or once every few seconds, for periods extending as well from a few seconds to a few minutes. The basic difference with actuated control systems is that newly developed plans generally become the new baseline for the next optimization. There is no reliance on background predefined plans.

2.2. TSP Priority Objectives

As indicated in the introduction, signal timing alterations are generally provided within TSP systems with the objective to reduce the number of stops and/or duration of stops incurred by buses due to the operation of traffic signals. While the reduction of delays and travel times has historically been the primary goal behind the implementation of early TSP systems, such systems are nowadays being implemented for a variety of reasons that may include one or more of the following:

- Reduction of transit delays at signalized intersections.
- Reduction of overall transit travel times.
- Provision of improved schedule adherence, service reliability and regularity.
- Reduction of vehicular traffic demand by providing enhanced transit service.
- Reduction of vehicle emissions and fuel consumption.

While priority can be given without consideration for cross-street traffic, such an approach is usually not desirable, nor easily justified by traffic engineers, due to its potentially significant negative impacts on overall network traffic performance. The challenge with the implementation of TSP systems thus generally becomes to give priority to transit vehicles with appropriate consideration for potential negative impacts on other traffic.

2.3. Transit Signal Priority System Components

There are three major components to a TSP system:

- Transit vehicle detection system.
- Communication system.
- Traffic signal control system.

The schematic in Figure 2.1 illustrates the relationship between these components. The transit vehicle detection is the component responsible for alerting the signal control system of an approaching bus. Upon detection of a vehicle, information regarding the position of the vehicle is sent to the traffic signal controller in charge of the intersection, either through a local or global communication system. After estimating the most likely arrival time of the approaching bus, the signal controller then determines whether signal timing adjustments may benefit the approaching bus and to implement these adjustments if found beneficial.

The following subsections present a more detail overview of the role of each component and of their interactions with other system components.

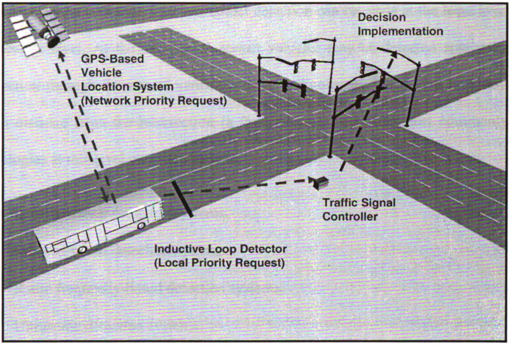


Figure 2.1 Schematic sketch shows TSP system components.

2.3.1. Transit Vehicle Detection System

The provision of priority to transit vehicles at signalized intersections relies on knowing that a vehicle is approaching an intersection or has already reached the intersection stop line. Once a bus has been detected, it is then the task of the signal controller to make a decision regarding the appropriate signal timing alterations to implement to reduce transit wait times while not overly negatively impacting other traffic.

The vehicle detection system is responsible for initiating requests for priority based on predefined criteria. The criteria used for initiating requests may be static in nature, such as automatically generating of requests for all buses on certain routes when they arrive at a certain distance from an intersection, or dynamic, such as the generating requests only for buses that are behind schedule by more than a certain amount of time or buses that have a certain number of persons onboard.

Depending on the approach selected, transit detection can be made at the local level or at the network level. At the local level, transit vehicles may be detected approaching an intersection through the use of various sensing technologies that do not look beyond a certain distance from the intersection (a few hundred feet). The most commonly used technologies at this level currently include:

- Inductive loop detection systems.
- Infrared-based detection systems.
- Radio frequency-based detection systems.
- Ultrasonic detection systems.
- High-intensity light detection systems.

Detection at the network level is accomplished through an automatic vehicle location (AVL) system, where vehicles communicate their position to a centralized transit or traffic management center through use of the Global Positioning System (GPS) or other vehicle location method. In this case, priority requests are generated based on the reported position of each vehicle and the expected time at which the bus is forecasted to reach the prioritized intersection. Appropriate TSP actions would then be selected either at the central traffic management office and passed down to the local signal controller or selected directly by the local controller after reception of information from the central management center that a priority action has been requested and can be granted.

The following sub-sections provide more detailed explanations on how selective transit vehicle detection is accomplished with each of the technologies mentioned above.

Inductive Loop Detection System

Inductive loops operate on the principle of inductance, i.e., the property of a wire or circuit element to "induce" electric currents in isolated but adjacent conductive media. A detector consists of an insulated electrical wire, placed on or below the road surface, attached to a signal amplifier, a power source, and other electronics. Driving an alternating current through the wire generates an electromagnetic field around the loop. Any metallic mass passing through the field will absorb electromagnetic energy and decrease the inductance and frequency of the loop. For most conventional installations, a change in inductance or frequency beyond a preset threshold will then indicate that a vehicle has been detected.

Many factors determine loop inductance, including wire size, wire length, the number of turns, lead length, and insulation. However, detector manufacturers are increasingly producing detectors having the capability of outputting a vehicle's inductive signature. When compared to a set of templates representing typical vehicle inductive signatures, the extracted profile can be used for classifying passing vehicles and selectively detecting specific vehicle types, such as transit vehicles. Older technologies generally relied on the use of an onboard transponder that would emit a unique signal that would be recognized by the system monitoring the inductive loop [5].

One of the primary disadvantages of such a detection system is its immobility. Since loop detectors are embedded into the pavement, they cannot be moved easily. In particular, any change in the location of detectors would imply making additional cuts into the pavement and possibly installing new power and communication cables in the ground at some costs.

Radio Frequency-Based Detection System

Radio frequency based system uses a low frequency radio signal to establish communication between a transit vehicle and a local signal controller. Transit vehicles are equipped with a radio beacon transmitter, while the local controller is equipped with a receiver. When a transit vehicle approaches the intersection, the controller receives the radio signal emitted by the vehicle. The signal is then decoded to verify the identity of the bus, its travel direction and assess priority needs. Some of the disadvantages of this technology include the potential for interferences from surrounding power lines, as well as a need to control the signal range to avoid influencing nearby signals and systems operating with similar technology.

Ultrasonic Detection System

Ultrasonic detection systems use sound energy to identify transit vehicles. In order for this system to work, a set of phased-array microphones is mounted on the traffic signal mast arm and set in a directional position parallel to the expected source of sound signal. A vehicle is detected when the microphones pick up changes in sound levels that are beyond a certain threshold. In this case, selective detection of transit vehicles is based on the recognition of specific sound patterns associated with these vehicles. However, one of the main potential shortcomings for this detection technology is noise interference from the surrounding environment or weather conditions (wind noise, rain noise).

Infrared-Based Detection System

Infrared detection relies on two components: a transmitter that is mounted on transit vehicles, and a receiver that is generally mounted on a traffic signal support. Detection occurs when an infrared signal emitted by the transmitter is detected by the traffic signal receiver. Upon reception of an infrared signal, the phase selector then interprets the data encoded in the signal to identify the approaching vehicle, assess the distance at which the vehicle is located, and the need for priority action. The main advantage of this system is that the detection distance can be easily adjusted. However, similar to ultrasonic detectors, infrared detectors are generally subject to weather and environment interferences that may reduce their detecting ability and accuracy.

High-Intensity Light Detection System

High-intensity light detecting systems, also known as strobe light systems, are currently one of the most commonly used detection systems for TSP [5, 6]. Similar to infrared systems, this technology bases the detection of vehicles on the reception of a light signal emitted from an emitter mounted on a transit vehicle. The light signal is modulated in such a way as to provide a unique light signature different from other surrounding light sources (vehicle headlights in particular), Upon reception of a signal, a priority request is then sent to the signal controller, which will then decide whether a priority should be granted or not. Similar to other detector systems, strobe light detectors are subject to weather and environmental interference that may reduce their ability to detect an approaching vehicle.

GPS-Based AVL Systems

The Global Positioning System (GPS) is a satellite-based navigation system that has initially been developed by the U.S. Department of Defense for military purposes. The system was gradually implemented throughout the 1980s and made available to the public in the early 1990s following its completion and recognition of its potential importance in civilian applications [7].

The system is based on a network of 24 satellites that are used for determining the position of objects anywhere on the globe. Objects can be tracked by the system if they are equipped with an appropriate receiver allowing them to communicate with satellites that are in line-of-sight from their location. Communication with at least three satellites is required for establishing position. Position is determined by triangulation methods using the measured distance between the object and each satellite. Since distance is measured by assessing the time spent by a signal to travel between the two objects, any interference in the path of the signal may result in decreased measurement accuracy. Examples of interferences include presence of particles in the atmosphere and reflected signals from tall buildings.

Since there is no subscription for the system, the only costs for using the system are mainly associated with the need to purchase GPS receivers and equipment to transmit the assessed position to a dispatch center. There is in this case no need to deploy an extensive field infrastructure. Agencies wishing to use a system providing increased position and speed measurement accuracy can still choose to subscribe to a Differential GPS (DGPS) service. DGPS services use base stations located at known positions for calculating signal correction factors when calculating the time required by signals to travel between satellites and receivers. These correction factors are then broadcast and used by DGPS receivers to increase the accuracy of position measurements.

Because of low initial deployment costs and the ability to know at any time the location of buses, transit agencies are increasingly relying on the use of GPS to track the progression of their fleet of vehicles while on the road. Consequently, increasing efforts have been made to use the information provided by the resulting vehicle location system to determine appropriate priority actions at signalized intersections.

2.3.2. Communication System

The communication system is responsible for linking the detection system to the traffic signal control system. Its purpose is to transfer detection information to the control system. It also allows communication from management center to local controllers. Depending on the size of the communication system, it is important to specify the system characteristics that are required for the provision of efficient TSP, such as system capacity and speed. As an example, in an earlier TSP deployment within a SCOOT real-time traffic signal control system, it was decided that TSP actions should be provided locally, by the signal controllers themselves, and not by the central computer that is in charge of SCOOT signal operations, on the simple grounds that communication of information from the local intersection to the central computer and then back to the local controller could take as much as four seconds. The rationale for this decision was based on the assessment that a four-second lag was too long to allow efficient signal timing adjustment to be calculated, particularly when considering that a bus may or may not be traveling and changing position within this interval.

2.3.3. Traffic Signal Control System

The traffic signal control system is responsible for acting on the priority requests and making any applicable changes to the signal indications, via a centralized or local traffic signal controller, or a combination of both. In some systems, the local traffic signal controllers may be able to perform this function completely. In other systems, a centralized control system may instead be put in charge of granting priority requests. In this case, the local controller would only act an executor of the decisions made by the central traffic control system. In another variation, a centralized traffic control center may be instructed to determine a set of constraints that the local controller should follow, but leave the final decision as to which priority decision to implement to be made by the local controller based on the established constraints.

Depending on the system setup, the traffic signal control system may or may not make changes to the signal indications following the reception of a priority request. For example, if a local policy limits the number of priority activations to one per cycle, a second priority request received by the traffic signal control system would not result in further changes to the signal indications. The traffic signal control system is also responsible for ensuring that higher priority requests (e.g. emergency/railroad preemption) override other requests in order of priority.

2.4. Priority Strategies

There are generally three classes of TSP strategies;

- Passive strategies.
- Active strategies.
- Real-time strategies.

Passive priority strategies give priority to transit vehicles without the need for vehicle detection. Conversely, active priority strategies provide priority only after a transit vehicle has been detected and priority conditions are met. Real-time strategies finally include strategies in which both the transit and traffic needs are determined simultaneously and used in on-line signal timing optimizations. The selection of an appropriate strategy to implement depends on the characteristics of the transportation network, objectives of the transit agency, cost considerations, and factors associated with the performance of the traffic signal controller, as outlined in the paragraphs that follow.

2.4.1. Passive Priority Systems

As indicated above, passive priority strategies do not require a transit detection system. As such, these strategies can be described as static as they operate continuously regardless of whether a transit vehicle is present or not. Passive strategies mainly consist of signal timing modifications generally favoring transit vehicles based on historical traffic performance, but might also include geometric or infrastructure enhancements. Examples of passive strategies include:

- Adjusting signal settings, such as cycle length, green times and offsets in a way that favors transit movements.
- Transit-based coordination: Replacement of traditional traffic-based traffic signal coordination schemes by new schemes tailored upon transit needs.
- Increasing priority phase split: Arbitrary increase in the duration of the green phase serving transit vehicles to increase the probability that a transit vehicle would reach an intersection during a green signal.
- Phase splitting: Splitting priority phases into multiple phases and repeating these phases within the signal cycle to reduce the amount of time that a vehicle has to wait at a red signal.
- Queue jumps: Provision of high occupancy lanes or transit lanes at intersections to allow buses to bypass queues.

Passive priority strategies are considered efficient form of TSP in situations where transit frequencies are high and traffic volumes are low. In such a situation, the high transit service frequency will tend to produce bus arrivals in almost all cycles or every other cycle. In addition, the low traffic volumes will tend to minimize impacts on general traffic in cycles in which a transit vehicle is not present and in which prioritized timing is not theoretically required. Passive priority strategies may also be efficient at intersections where transit operations are predictable, as the resulting regularity of transit priority needs will then facilitate the establishment of permanent signal timing, signal phasing or road geometry alterations.

The main advantages of passive priority strategies are ease of implementation and low cost, mainly because they do not require a detection system. However, disadvantages may include increased delay to side-street traffic, excessive allocation of green time to priority movement, and dissatisfaction from motorists on non-priority movements experiencing increased stops and delay, particularly from motorists reaching the intersection in cycles in which no transit vehicles are present. Furthermore, if the transit headways are large, as is generally the case in small/medium size cities, these strategies may induce unnecessary delay to the entire system, since the prioritized signal timing will keep running even when buses are not there.

2.4.2. Active Priority Systems

Active priority strategies are dynamic signal timing enhancements in which the signal timings are temporarily modified upon the detection of an approaching transit vehicle. These strategies provides for an efficient operation by implementing signal changes only when needed and possible.

Active priority strategies are classified into two types: unconditional and conditional. Unconditional priority simply considers whether a bus has been detected or not. As such, this type of priority offers considerable potential for express transit during off-peak hours, especially when there are no limits on green extension or red truncation length. Conditional priority provides an enhancement by awarding priority to a detected bus only when specific conditions are met, such as the number of passengers onboard a vehicle, the extent to which a bus is on or behind schedule, the time since the last priority was awarded, or potential delays on other traffic.

Whether conditional or unconditional strategies are considered, the following signal timing alterations are considered to provide preferential treatment to buses:

- Early green (red truncation).
- Green extension.
- Phase rotation.
- Actuated transit phase (phase insertion).

Early Green (Red Truncation)

The early green strategy, also known as red truncation (RT), shortens the red time of the preceding phases to expedite the return of the green signal serving the movement on which a prioritized vehicle has been detected. This strategy is applied when the prioritized vehicle is expected to arrive at an intersection near the end of a red signal. Figure 2.2 illustrates as an example the changes occurring to a signal timing sequence when a transit vehicle is expected to arrive a few second before the priority phase starts. In this case, the traffic signal controller truncates the non-priority phase to allow the transit vehicles to leave the intersection a few seconds earlier than its regular time. This priority action is subject to minimum green requirements from other phases, as no green phase is usually allowed to be shortened beyond a certain minimum, such as the time required by pedestrians to cross an approach or a minimum queue serving time.

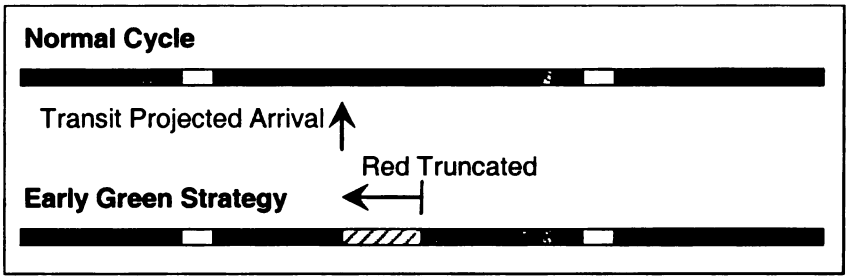


Figure 2.2 Early green strategy

Green Extension (Phase Extension)

The green extension (GE) strategy extends the green time allocated to a phase serving an approaching bus. This strategy is only applied when an approaching bus is scheduled to arrive a few seconds after the end of the green. Green extension is one of the most effective forms of priority action since an extension not only does not require additional clearance intervals and but also allows a transit vehicle to avoid a relatively long delay corresponding to an entire red duration. Figure 2.3 explains the concept of green extension, and what change would occur within a signal cycle when a priority is requested.

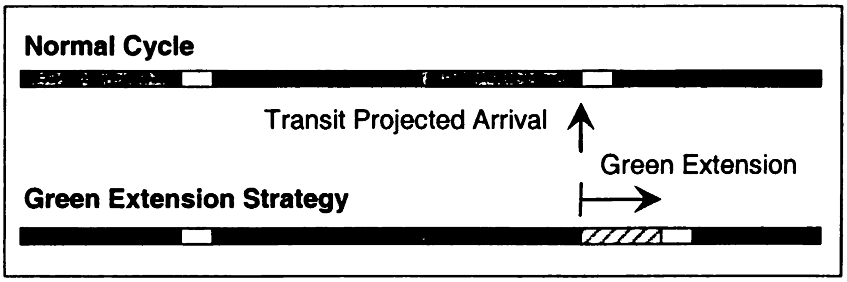


Figure 2.3 Green extension strategy

As is often the case, early green and green extension strategies may be applied together to maximize the time within the signal cycle in which transit would be eligible for priority. Depending on the system, early greens and green extensions can be offered within a fixed cycle framework, i.e., without modifying the cycle length, or within a variable cycle framework. A variable cycle framework provides the most flexibility as there is no need to compensate phase extensions by reducing the duration of other phases. However, since the use of variable cycles can disrupt traffic progression patterns in coordinated signal systems, TSP systems implemented along urban coordinated arterials have historically only be developed to operate within a fixed cycle setup. Nevertheless, as a result of the development of more complex signal controllers, some arterial priority systems are now starting to appear in which the cycle may be temporarily extended or shorted at individual intersections to accommodate priority requests. In most cases, an extension or shortening of the signal cycle is compensated by an opposite action in the next cycle to allow the signal system to quickly reestablish progression patterns.

Phase Rotation

As illustrated in Figure 2.4, the order of signal phases can also be "rotated" to accommodate approaching transit vehicles. In this case, a northbound left turn phase may follow a general northbound through phase. A northbound left turning bus requesting priority arriving before the start of the green phase serving through traffic could then request that the left turn phase be initiated earlier. With the phase rotation concept, the left turn phase could be instead displayed as a leading phase before the through phase instead of as a lagging phase in order to expedite the passage of the transit vehicle. When compared to the early green and green extension options, the phase rotation would have the advantage of maintaining the amount of green initially dedicated to each traffic movement. However, implementation of such a strategy is often not recommended as it might confuse drivers and lead to increased crash risks.

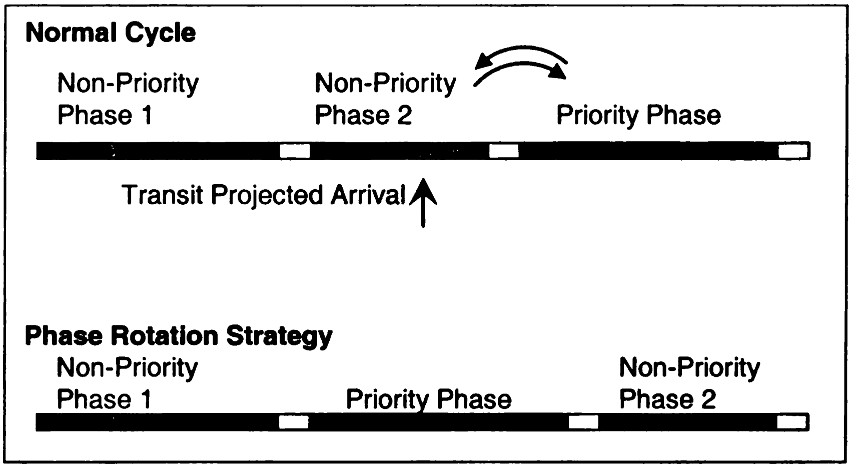


Figure 2.4 Phase rotation strategy

Actuated Transit Phases (Phase Insertion)

Actuated transit phases are only displayed when a transit vehicle is detected approaching an intersection. As shown in Figure 2.5, an example would be an exclusive left turn phase for transit vehicles. The left turn phase would only be displayed within the signal cycle when a transit vehicle is requesting priority of passage. A potential problem is the fact that additional lost time might be added into the cycle due to the need to add another interval. Another element to consider is the possibility for buses to benefit from the displayed phases. The potential for blockage by other traffic has led to the application of this priority action only at intersections where buses travel on exclusive lanes or where queue jumps have been added to allow a transit vehicle to bypass a queue of vehicles.

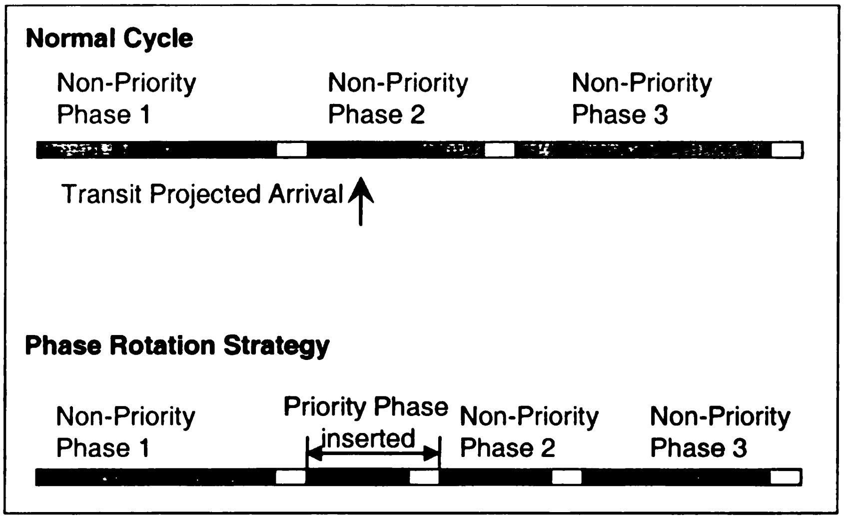


Figure 2.5 Phase insertion strategy

2.4.3. Adaptive (Real-Time) Systems

Adaptive/real-time TSP strategies provide priority while simultaneously trying to optimize traffic movements. Optimization is generally conducted using performance criterion linked to general traffic operations. The selected criterion may include person delay, transit delay, vehicle delay, and/or a combination of these criteria. Adaptive strategies continuously seek to optimize the active timing plan based on real-time, observed data. As such, they require early detection of both transit vehicles and traffic to obtain advanced information about short-near-term vehicle arrivals. Adaptive systems also often require the ability to update the transit vehicle's arrival time, which can vary due to the number of stops and traffic conditions. The updated arrival time can then be fed back into the process of adjusting the signal timings.

Over the years, there have been many attempts to incorporate TSP within real-time traffic signal control systems. Below is a summary how SCATS, SCOOT, OPAC, UTOPIA and RHODES were designed or are proposing to handle TSP within a real-time operational optimization context.

SCATS

The SCATS (Sydney Coordinated Adaptive Traffic System) traffic adaptive signal control system was originally developed for the New South Wales Roads and Traffic Authority for application in Sydney and other Australian cities. Over the years, this system has been implemented in over 80 cities worldwide, including Australia, New Zealand, and East Asia. In the U.S, it has been deployed in Oakland County, Michigan, Delaware and Minnesota. Its main operational objective is to adjust cycle time, splits and offsets in response to real-time traffic demand with the goal to minimize overall stops and delays [5].

The hierarchical control architecture in SCATS consists of two levels: a strategic level and a tactical level. At the strategic level, a regional computer coordinates the timing of signals in a network or subsystem consisting of one to ten signals. Subsystems can link together, to form larger systems operating on a common cycle time, where each subsystem has only one critical intersection. At the tactical control level, signal timing decisions are distributed among the individual intersections. Split plans are selected by individual local controllers within the constraints imposed by the regional computer's strategic control. Each local controller bases its signal control decisions on data collected from sensors monitoring how traffic utilizes the available green time. The objective here is to maximize green time utilization by considering terminating a phase when the demand for the green drops below a certain threshold.

With respect to TSP, SCATS is capable of providing both passive and active priority. Passive priority treatments can be granted to transit vehicles by allocating a favorable bias to the links traveled by these vehicles in the signal coordination plans considered by the regional computer. On the other hand, active priority requests can be granted by local controllers as part of their tactical control decisions. Upon detection of an approaching bus, priority phases can be called to either clear the queue ahead of the vehicle or to provide a temporary phase extension. Priority requests can be granted or denied depending on the time of day, tidal flow determination based on traffic flows or on how congested the intersection is. Following the granting of a priority request, the system would automatically compensate for any negative impacts on traffic by reacting to demands for longer green time by vehicular traffic on the approach for which the green time may have been reduced.

Because SCATS' stop line traffic detector cannot provide advanced information about approaching transit vehicles, TSP is only possible if additional detectors located at some distance upstream of the intersection are installed.

SCOOT

SCOOT (Split Cycle Offset Optimization Technique) is a centralized traffic signal control system that was designed to automatically respond to fluctuations in traffic demand through the use of detectors located at some upstream distance from the intersection. The system uses the detectors to build a profile of traffic demand in the next signal cycle and uses this cyclic flow profile to determine whether it may be beneficial to change by a few seconds the time at which individual phases are scheduled to be terminated. Signal control decisions within SCOOT are taken not by individual signal controllers, but by a central computer that is in charge of optimizing signal timings at all intersections within a given region. This system is the predominant real-time traffic signal control system in operation today, with installations throughout Europe, Asia and North America [8, 9].

Similar to SCATS, SCOOT can provide either passive or active preferential treatments. Passive treatments can be implemented by introducing biases for the links traveled by transit vehicles or imposing fixed offsets between signalized intersections based on average transit travel times. Active priority treatments can also be offered. Such treatments are granted by local controllers as an override of signal timing decisions made by the central computer, although some systems may have the central computer taking an active role in the granting of priority requests [10]. Upon detection of an approaching bus, a signal control first checks whether it has been granted permission to override network control decisions. Permissions to override network control may for instance be denied if traffic demands are too high and the intersection is operating near capacity. If overrides are possible, the signal controller will then assess whether a green extension or early green recall can be offered, within the limits set by the central computer, to provide preferential treatment to the approaching bus.

To avoid bus arrival prediction problems caused by variable dwell times, it is generally required within SCOOT systems that bus detection occurs downstream of bus stops. This means that when SCOOT system is implemented, bus stops must be located on the far-side of an intersection or sufficient upstream of an intersection approach to provide advance warning of an approaching bus [9].

OPAC

Optimization Policies for Adaptive Control (OPAC) is a computational strategy for real-time demand responsive traffic signal control [11] that was initially developed to control isolated intersections but that has been expanded to network control. This system uses a dynamic programming approach to determine the combination of signal switching decisions that maximizes overall effectiveness or minimizes overall disutility. The optimization algorithm tries to find a short-term local signal control solution that would be close to near-optimal operation based on predicted and actual vehicle arrivals. OPAC progressively selects a signal pattern from among a number of possible solutions for each intersection individually. The selected patterns are recalculated every few seconds based on updated traffic data [5]. TSP is implemented within OPAC by modifying the algorithm so that signal settings are optimized based on number of passengers instead of vehicles flow [12].

UTOPIA

UTOPIA is an urban traffic control system that was first developed and implemented in Italy in 1985 with the objective of integrating real-time traffic control with transit priority needs. The system places a strong emphasis on the decentralization of signal timing optimization. It is a two-level control system in which signal controllers play a dominant role in the development of signal timing strategies. The controllers determine the signal settings to be implemented at each intersection according to local traffic demands, transit priority needs, and coordination needs with adjacent intersections. The central computer adds robustness and stability to the control strategies developed by the local controllers by imposing timing constraints based on network-wide coordination objectives and a network-based reference timing plan.

As indicated above, transit priority is primarily considered at the intersection level. In the simplest setup, priority is granted on a simple call basis by local controllers, with no direct consideration for the potential disruptions to other traffic. A more complex implementation in the original system developed for the city of Turin in Italy had a more complex bus arrival prediction system allowing individual signal controllers to get up to 8-min advance warning of an upcoming bus arrival. Each time a transit vehicle is detected, a message is sent to the central computer asking it to predict the arrival times of the vehicle at successive downstream intersections for the next 8 minutes. These times are updated each time a bus passes another detector [11, 13], and the updated information is used to adjust the signal timings to better match the projected bus arrival.

RHODES

RHODES was developed based on the concept of predicting arrival of vehicles on all approaches to each intersection and using the resulting predictions to explicitly optimize phase durations to produce a given optimal performance according to a user-defined measure, such as minimizing average delay or number or stops. Similar to UTOPIA, RHODES features a hierarchical control structure. At the highest level, the system uses a dynamic network loading model that captures general traffic characteristics variation. The traffic characteristics considered at this level pertain to the network geometry, travel demand between origins and destinations, and typical route selections of travelers. Variations in these characteristics provide RHODES with general allocations of green times for each different demand pattern. The resulting green time decisions are then updated at the middle level of the hierarchy, referred to as "network flow control." At this level, traffic flow characteristics are measured in terms of platoons of vehicles and their speeds. Signal timings generated at the upper level are adjusted here to adequately service expected platoon arrivals. The resulting timings are then passed down to the third level, the "intersection control" level, which will then select the appropriate phase change epochs based on observed and predicted arrivals of individual vehicles at each intersection.

TSP is implemented in RHODES in two different ways referred to as the weighted buses approach and the phase constrained approach. In the weighted buses approach, transit vehicles are converted to a certain number of passenger cars equivalent based on their lateness and their passenger load. In the phase constrained approach, the TSP is implemented by scheduling the green time for transit vehicles according to the time each transit vehicle is expected to arrive at an intersection [14-16].

2.5. Signal Recovery

Although signal recovery is not a particular TSP strategy, it is important to mention its role in system performance. TSP actions do not always end when a transit vehicle passes through an intersection and the priority phasing is terminated. For equity reasons, many traffic engineers will require that signal controllers follow up on the priority actions with a recovery operation, in which the signal controller would transition back to normal signal operation. The idea behind such transition is to compensate for time lost by the non-priority movements or maintain coordination along arterials. Examples of recovery actions commonly considered include:

- Denial of priority for *n* cycles following the granting of a priority request.
- Provision of extended green time to the non-prioritized traffic that was penalized by the priority action.

The implementation of signal recovery plans is seen by traffic engineers as playing a vital role in mitigating impacts on cross streets and helping to maintain coordination along urban corridors. However, it must also be recognized that the implementation of recovery methods can significantly limit the ability of a priority system to respond to the needs of all buses requesting priority. For this reason, signal recovery plans must be carefully designed if their use is warranted.

CHAPTER III.

PROBLEM DEFINITION

Numerous studies have already demonstrated that transit vehicles benefit from traffic signal priority systems [17-20]. However, many of these studies have also shown a potential for increased traffic delays as a result of the need to temporarily shorten certain green phases or extend the signal cycle to accommodate approaching buses. These delay increases can be particularly significant at busy intersections where most of the available green time may already be required to serve existing traffic. Another concerning element is the potential loss of progression patterns along coordinated arterials, which may result in delay increases at a number of successive intersections.

Due to the potential for negative impacts, systems providing effective preferential treatments to buses while keeping traffic impacts at a minimum are generally sought. This means developing systems that consider all potentially impacting parameters when making signal timing alteration decisions. One parameter to consider is uncertainty in predicting transit travel times. The importance of this parameter is linked to the uncertainty regarding the time a bus will need to board and alight passengers. If dwell times were perfectly predictable, stop line bus arrival times could always be accurately predicted. However, dwell times generally exhibit some variability, which creates uncertainty in arrival predictions. In turn, this may lead to the implementation of green extensions and early recalls that may not be used by the transit vehicles at the source of their implementation.

Based on observations from several U.S. cities, the Transit Capacity and Quality of Service Manual, or TCQSM [21], indicates that bus dwell times are characterized by a coefficient of variation ranging from 0.4 to 0.8. Among the impacting elements responsible for such a variation are:

- Passenger demand and loading, particularly the number of persons passing through the highest-volume door and the proportion of alighting and boarding passengers attempting to use the same door.
- Bus stop spacing, as fewer bus stops lead to increased demands at individual stops.
- Fare payment procedure, as payment with exact change can take up to an additional 1.5-s per passenger than payment with a pass.
- Vehicle type, as having to ascend or descend steps increases the time required to serve each passenger.
- In-vehicle circulation, as extra time may be needed to allow passengers to reach the door or the fare box when persons are standing within the vehicle.

- Other specific needs, such as time required for wheelchair loading.

Specific problematic elements covered in this dissertation and outlining challenges to be considered in the proposed research include:

- The problem of bus stop location.
- The problem of variable dwell times.
- TSP within coordinated signal systems.
- Simultaneous TSP requests.

3.1. Bus Stop Location

As the point of access to transit services, the bus stop is a critical element in a transit system's overall goal of providing timely, safe, and convenient transportation. On this note, TCRP Report No. 19 summarized several concerns of interest to transit service users and riders regarding the design and location of transit stops [22]. These concerns are:

- *Transit system performance:* Travel time for a bus trip has four components: the time it takes to walk to the bus stop, the wait time for the bus, the actual in-vehicle travel time, and the time to walk to the destination. Within this context, transit system performance is strongly affected by the location of bus stops and the frequency at which buses service each stop.

- *Traffic flow impacts:* Both the location and design of a bus stop affect the flow and movement of other vehicles. In particular, a well-designed bus stop can allow passengers to board and alight without the bus significantly impeding or delaying adjacent traffic.

- *Safety:* Safety is the freedom from danger and risk. In a transit environment, this concept includes an individual's relationship to buses and general traffic, and the relationship of a bus to other vehicles. Safety issues for pedestrians include nearness of a bench to the flow of traffic on a busy street, or the ability to safely cross a street to reach a bus stop. For buses, safe reentry into the flow of traffic is a prime example of operational safety concern. In general, different safety levels may be associated with different bus stop location and layout, often making this evaluation context-specific.

- *Security:* Security refers to an individual's feeling of well being. Security is affected by lighting at bus stops, bus stop visibility from the street and nearby land uses, and presence of hiding places at bus stops.

The above elements define functional and performance-related concerns for public transportation systems. Each must be addressed to achieve the goal of timely, safe, and convenient public transportation and to satisfy the needs of the service area. More importantly, as was outlined, each area of concern is influenced by the selected location and layout of bus stops.

For TSP systems, a particularly important element is the bus stop location relative to the intersection. Bus stop locations can be categorized as far-side, near-side, and midblock. Figure 3.1 illustrates the physical meaning of each of these categorizations.

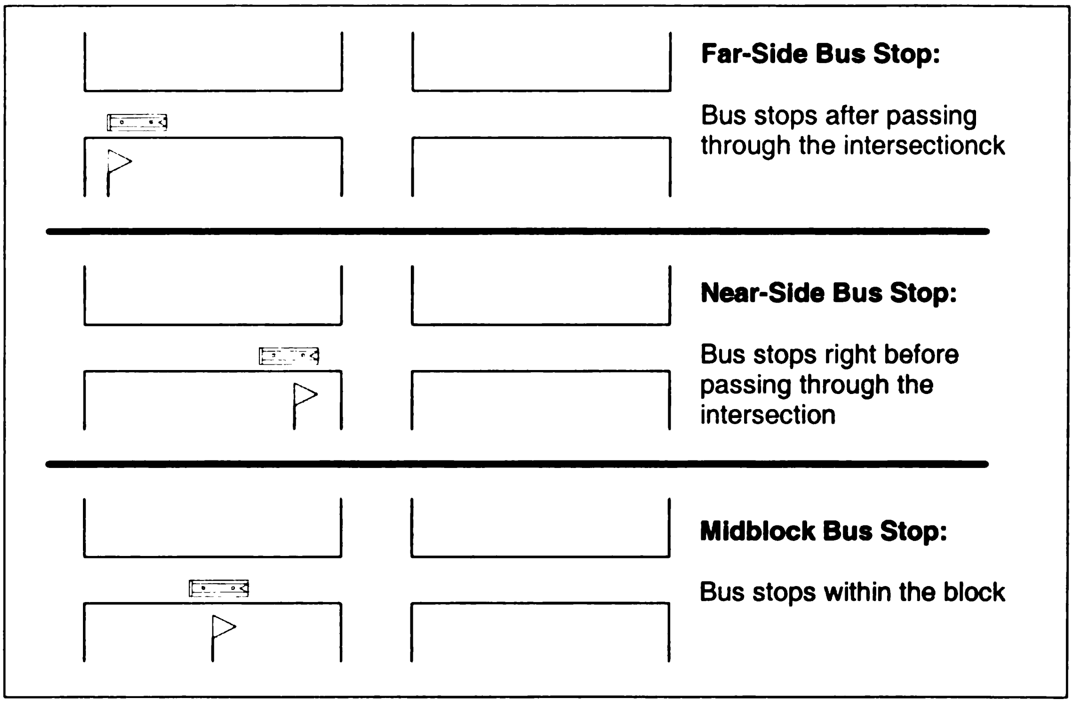


Figure 3.1 Schematic sketch for bus stops location

Each type of bus stops has its own advantages and disadvantages, which often lead individual agencies to favor one type over the other. The TCRP report [22] suggests the consideration of the following factors when selecting which type to implement at a particular location:

- Adjacent land use and activities.
- Bus route (for example, is bus turning at the intersection).
- Bus signal priority (e.g., extended green suggests far side placement).
- Impact on intersection operations.
- Intersecting transit routes.
- Intersection geometry.
- Parking restrictions and requirements.
- Passenger origins and destinations.

- Pedestrian access, including accessibility for handicap/wheelchair patrons.
- Physical roadside constraints (trees, poles, driveways, etc.).
- Potential patronage.
- Presence of bus bypass lane.
- Traffic control devices.

Consideration of the above elements may result in a bus stop placement on the near-side or far-side of an intersection, or at some distance upstream. For TSP, the most problematic placement is a near-side stop, where the bus would stop at or very close to the intersection stop line. This is problematic as such a placement creates difficulty in predicting the exact moment at which a bus may need to enter the intersection due to the uncertainty in knowing the exact amount of time that a bus will need to remain at the stop to board and alight passengers. Since the bus stop is close to the stop line, there is very little time to respond to a bus leaving the stop. Prediction models must then be used to assess the most likely bus stop line arrival times. As will be detailed later, this creates problems for the signal controller in correctly assessing the priority needs and the actions that need to be undertaken. While moving the bus stop further upstream may allow bus detection downstream of the stop and remove the problem of variable dwell time, such an upstream location introduces new problems in the form of travel time uncertainties in mixed traffic, where traffic interactions may delay the progression of a bus beyond what may be expected.

As can be assessed, there is no clear ideal bus stop location. This assessment is further emphasized by conclusions from TCRP Report 19, shown in Table 3.1, which summarize the advantages and disadvantages associated with each bus stop type. Because of the difficulty of accurately predicting bus arrivals with near-side stops, current practice with respect to TSP systems is to prefer midblock or far-side stops, to allow bus detection downstream of bus stops. Many TSP equipment providers even go as far as suggesting moving near-side bus stops. However, consideration of cases with near-side stops should not be ignored, as many transit agencies may prefer these stops for various reasons (such as facilitating transfer, reducing street crossings) or may physically be unable to move near-side stops to a "more suitable" alternate location. Moreover, the deployment of TSP should be flexible enough to take into account the different types of bus stops. Therefore, the need for a comprehensive TSP model that is insensitive to the bus stop type is important.

Table 3.1 Comparative Analysis of Bus Stop Locations

	Advantage	Disadvantage
Far-side Bus Stop	• Minimize conflicts between right turning vehicles and buses. • Provide additional right turn capacity by making curb lane available for traffic. • Minimize sight distance problems on approaches to intersections. • Encourage pedestrians to cross behind the bus. • Creates shorter deceleration distance for buses since the bus can use the intersection to decelerate • Resulting in bus drivers being able to take advantage of the gaps between traffic flows that are created at signalized intersections.	• May result in the intersection being blocked during peak periods by stopping buses. • May obscure sight distance for crossing vehicles. • May increase sight distance problems for crossing pedestrians. • Can cause a bus to stop far side after stopping for a red light, which interferes with both bus operations and all other traffic. • May increase number of rear-end accidents since drivers do not expect buses to stop again after stopping at a red light. • Could result in traffic queued into intersection when a bus is stopped in travel lane.
Near-side Bus Stop	• Minimize interferences when traffic is heavy on the far side of the intersection • Allow passengers to access buses closest to the crosswalk. • Results in the width of the intersection being available for the driver to pull away from curb. • Eliminates the potential of double stopping • Allow passengers to broad and alight while the bus is stopped at the red light • Provides drivers with the opportunity to look for oncoming traffic, including other buses with potential passengers.	• Increases conflicts with right-turning vehicles • May result in stopped buses obscuring curbside traffic control devices and crossing pedestrians • May cause sight distance to be obscured for cross vehicles stopped to the right of the bus • May block the through lane during peak period with queuing buses • Increases sight distance problems for crossing pedestrians.
Mid-Block Bus Stop	• Minimizes sight distance problems for vehicles and pedestrians. • May result in passenger waiting areas experiencing less pedestrian congestion.	• Require additional distance for no-parking restrictions. • Encourage patrons to cross street at midblock (jaywalking). • Increases walking distance for patrons crossing at intersections.

Source: Transit Capacity and Quality of Service Manual, 2nd Edition. TCRP Report 100, ed. T.C.R. Program. 2003, Washington, D.C.: TRB, National Research Council

3.2. Variability of Transit Dwell Times

As indicated earlier, it is often recommended that priority systems be only implemented at intersections with far-side stops or where stops are sufficiently upstream to allow vehicle detection downstream of the stop and to reduce uncertainties in predicting intersection bus arrivals created by variable dwell times [19, 23-26]. While such recommendation may satisfy the operational practices of many agencies, stop relocation may not always be possible.

Dwell time is defined as the time a bus is immobilized for passengers boarding and alighting. This only includes the time during which doors are opened. Another delay related to dwelling activities is the bus stop clearance time. For buses stopping in a traffic lane, this is the time needed to start up and leave the stop after the doors have closed [21]. For buses stopping in a bus stop, the time spent waiting for a gap of sufficient length in the adjacent traffic is also included. For signal priority systems, dwell and clearance times are often conveniently combined into a single operational parameter. Since signal timing adjustments are primarily concerned with the total delay incurred by vehicles at bus stops, combining both parameters allow system operators to consider a single parameter reflecting what they are primarily concerned with.

To illustrate issues associated with dwell time variability consider an approach with a near-side stop on which buses are detected when 20-s away from the bus stop and stop line. If buses always dwell for 30-s at the stop, signal timing alterations based on a projected stop line arrival 50-s after detection would always reflect reality. However, as illustrated by the example of Figure 3.2, dwell times are rarely constant. Here, while a 10-s dwell time would adequately characterize average bus operations, it would not reflect the fact that individual stop durations vary between 6 and 16 seconds and that about 58% of all buses were observed to not stop at all, thus creating a 16-s overall window of potential arrival times at the downstream intersection.

While the above example does not consider the effects of clearance times, as buses at the observed stop board and alighted passengers while stopped in a traffic lane, situations in which buses would have to seek gaps in adjacent traffic should exhibit even greater variability. According to the TCQSM, a bus needs 10-s to start up and clear a stop after closing its doors. For buses stopping in a bus stop, field observations further indicate that the average reentry delay can vary between 1 and 15 seconds, depending on the traffic volume on the adjacent lane. Consideration of these statistics would thus clearly add to the potential variability of total bus dwell times as considered by TSP systems.

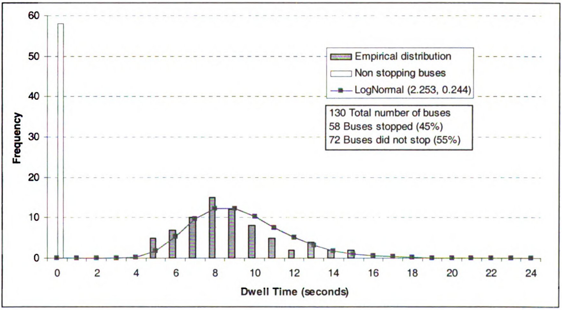


Figure 3.2 Observed dwell time distribution at a bus stop in Kitchener, Ontario, Canada.

While several studies have examined the performance of transit signal priority systems under various geometric, traffic, and transit conditions [17-20], these studies have primarily focused on intersections with far-side stops or where detection can occur downstream of the stop. Although it is possible to use an average dwell time, this might have a significant impact on evaluating the proposed traffic signal controller [27, 28], and further need is required to predict dwell times at bus stops.

Among the few studies that have considered near-side stops, a significant focus has been on investigating bus arrival prediction techniques. Research examples include the use of linear and nonlinear regression models to predict dwell times based on the number of boarding and alighting passengers, the average headway between buses, schedule adherence, the number of doors on the vehicle, the fare collection method, and bus type [29, 30]. In other efforts, ANN-based models have also been proposed for predicting bus arrival times at intersections through improved consideration of historical link travel time and bus stops demand [31]. However, these models depend mainly on historical data to develop a prediction model, which will reduce their efficiency in real-time predictions, especially where a highly accurate and reliable dwell time prediction is required. Furthermore, the transit priority system and traffic control performance highly depends on the accuracy of dwell time prediction model, and its ability to go beyond data used to develop the prediction model [27, 28].

3.3. TSP Within Coordinated Traffic Signal Control Systems

Implementing TSP at an isolated intersection or at intersections that are part of a coordinated network presents different difficulty levels. At an isolated intersection, there is no consideration for the need to maintain traffic progression patterns with adjacent intersections. In such a case, there is great flexibility in selecting which signal timing alterations can be implemented. However, a majority of TSP applications that have been considered in recent years involve deployments along coordinated arterials. In these cases, some constraints must be considered when determining the best priority actions to implement in order to maintain decent or specific traffic progression patterns.

From a network point of view, there are generally two different approaches to providing TSP. In its simplest form, TSP is provided one intersection at a time, without coordination constraints. Following the detection of an approaching vehicle, the signal controller in charge of the intersection at the end of the approach would consider possibilities for providing priority to the approaching vehicle at the intersection under its control without factoring potential effects on traffic patterns or priority opportunities at intersections further downstream. Most TSP implementations are currently of this type [32]. A particular advantage of this approach is the removal of the need to track vehicles across multiple intersections, which greatly simplify system implementation, particularly if vehicles are not already equipped with automated vehicle location systems. Handling priority on an intersection-by-intersection basis also removes the need to develop procedures for predicting vehicle arrivals at successive intersections.

In a network setting, however, intersections are often linked together through the imposition of signal coordination constraints. In such systems, signal timings at one intersection are function of the timings at adjacent intersections and determined not necessarily to simply minimize stops and delay at the local intersection but also to minimize stops and delays across a succession of intersections through the provision of smooth traffic progression patterns. The primary difficulty for TSP systems in such a control environment is how to provide preferential treatments while maintaining signal coordination and effective traffic progression patterns.

TSP issues within coordinated systems are not only linked to a potential for reduced green time along main travel directions. Priority actions granted to buses traveling along an arterial with the main flow of traffic could for instance significantly alter established progression patterns by allowing traffic to be released earlier from an intersection. While this may reduce delays at the intersection in question, it may also result in vehicles reaching the following intersections earlier than expected. Since vehicles are reaching downstream intersection earlier than expected, they could be forced to stop again, causing additional delays at this intersection and potentially contributing to increasing queue lengths and side-street blockage.

The potential for altering traffic progression patterns introduces a need to consider operational constraints limiting the magnitude or type of signal timing alterations that can be implemented. As an example, some transit agencies limit for instance green extensions or early recalls to 5 seconds in order to minimize impacts on traffic progression patterns. While such a constraint may seem arbitrary, its imposition is often the results of an inability to quickly assess the traffic impacts that may arise at downstream intersections within a coordinated system following priority decisions taken at a particular intersection. In an ideal system, constraints to impose on priority actions should be tailored to the actual potentials for traffic impacts, being more or less relaxed depending on the situation.

3.4. Simultaneous TSP Requests

A particular difficulty with current TSP systems is what to do when two or more buses are approaching an intersection simultaneously. In most cases, TSP systems would grant priority on a first-come-first-serve basis. This is a carryover of past years, when technology did not allow tracking vehicles or obtaining real-time information about ridership or schedule adherence. Technological advances now allow controllers to consider these elements, and thus, to theoretically provide priority to the bus yielding the highest benefits, such as to the bus with the highest ridership or the bus that is the furthest behind schedule. However, for these criteria to be effectively considered, a system must exist to forecast needs over the next few seconds or cycle to allow optimum decisions to be made and avoid situations in which a switch to a more efficient action is prevented by constraints associated with earlier decisions.

3.5. Summary

This chapter focuses mainly on the issues that might happen when TSP is implemented. One of the issues is the location of bus stops. Placing the bus stops on the far-side of an intersection allow the controller to ignore variations in dwell time and to detect transit vehicles from a certain distance from the intersection. One potential problem is however the need to accurately predict bus travel time between the detection point and intersection stop line in mixed traffic. Placing the bus stops on the near-side of an intersection reduces the travel time variation but introduces the need to consider dwell time variations. Due to specific operational factors, transit agencies may be required to use both types of bus stops in order to reach their service objectives, such as having a bus stop to be a transfer point to other transit routes, or minimizing passenger's walking distance.

The uncertainty in predicting dwell times further affects the controller ability to predict when a transit vehicle is expected to enter and leave its target intersection. As outlined above, this issue is more serious at intersections with a near-side bus stop. Such an intersection will leave a limited timeframe to the controller to implement a successful TSP strategy.

Another potential problem regarding TSP implementation is the provision of preferential treatment in coordinated traffic signal control systems. While many traffic signal control systems are designed to consider network-based traffic needs and perform signal optimization by considering the collective traffic needs within a given area, TSP systems often only focus on providing preferential benefits to buses at individual intersections. This lack of coordination between transit and traffic operations needs may lead to decreased benefits from the implementation of TPS systems, either through increased traffic stops and delays as a result of loss of coordination or through the imposition of constraints unduly limiting signal timing alteration options for granting priority to buses [33].

Finally, the inability to handle simultaneous and near-simultaneous priority requests might prevent a TSP system from implementing truly optimal signal timing alterations. The particular significant of this problem is that as TSP systems are increasingly being deployed in urban areas, there is an increasing likelihood that the desire the priority multiple transit routes within a given area will lead to conflict situations at intersections crossed by multiple routes.

CHAPTER IV.

RESEARCH OBJECTIVES AND APPROACHES

This chapter briefly outlines the objectives at the base of the research project detailed in this dissertation. The overall objective of the research is to seek ways to improve TSP operations within coordinated arterial networks, more specifically with respect to the various technical issues presented in Chapter 3. This chapter highlights the objective of this research in details, introduces the optimization and prediction methods that will be utilized for the development of the proposed TSP controller. The need for the use of a microscopic simulation model to act as a virtual testbed for the development and evaluation of the proposed traffic signal control algorithm is also presented.

4.1. Research Objectives

The objective of this research is to be achieved through the development of Dynamic Signal Priority Optimization in Real-Time Traffic Control System, (D-SPORT), which is a traffic signal control that integrates TSP within the real-time signal control in a coordinated traffic network. The development of the D-SPORT controller is to overcome the challenges associated with implementing TSP in the coordinated system. These challenges include:

- Need for predicting bus intersection arrival times that is insensitive to the bus stop type that is used, particularly those with near-side stops or where bus dwell times may be highly variable. This need is driven by the fact that accurate information about the expected bus arrival time is a key to provide efficient TSP control.

- Need to minimize impacts on surrounding traffic performance of surrounding intersections when providing preferential treatment to buses.

- Need to efficiently handle situations in which conflicting priority requests may be received from more than one approach.

The D-SPORT controller consists of prediction and optimization processes. While the prediction process is used to provide the D-SPORT controller with the ability to keep track of the location of transit vehicles in the near future, the optimization process is used to find a network-wide signal timings to accommodate both transit and non-transit traffic throughout the network.

4.2. Research Approach

To develop an efficient TSP decision model that has the potential to address all of the problems discussed earlier, it is proposed to develop control algorithms to leverage novel emerging optimization techniques. For this task, genetic algorithms and artificial neural networks are used to develop a real-time traffic signal control and TSP integration model that has the capability to be efficiently applied in a variety of control situations. Application of these techniques not only requires developing a fundamental understanding of their operations, but also how a traffic signal control problem can be translated into an appropriate model suitable for each technique. This requires not only identifying the variables to consider, such as bus detection times, traffic volumes and expected bus dwell times, but also identifying how these variables can be used to generate efficient traffic signal timing control decisions.

4.2.1. Traffic Signal Optimization Methods

As indicated in the problem definition, this research aims to develop a traffic signal optimization methodology to find near-optimal signal timings at a coordinated network of signalized intersections. Optimization can be defined as the process by which a "best" solution to a problem under analysis can be found by means of application of scientific methods and tools [34]. The following subsections present a quick review of available optimization methods and a justification for the selection of the Genetic Algorithm approach as the traffic signal optimization tool for the development of the D-SPORT controller.

Calculus-Based Optimization

Calculus-based optimization relies on purely mathematical techniques to find the largest or the smallest value a function would take in a specified interval. This optimization method requires the development of mathematical formulation replicating as close as possible the real process that is to be optimized. As such, calculus-based optimizations are only valid as long as it can be certified that the mathematical model used in the optimization is adequately representing reality. This has been a significant problem in the application of such an optimization technique to traffic control problems, as traffic behavior is a highly complex and dynamic process that is difficult to represent by simple mathematical equations. Other limitations of this method include the assumption that the function used in the optimization should be continuously differentiable. Furthermore, this method is unable to handle the fact that the control variables are restricted to a set of control functions that are used to validate the control variables.

Dynamic Programming

Dynamic Programming (DP) deals with problems involving a finite number of stages but in which the system being optimized may be in several possible states. The general aim of DP is to determine an optimal policy that minimizes the total costs in all the stages. A problem is solved recursively by finding, at each decision point, an action that minimizes the sum of the current cost and the best future costs. A typical example of DP application is to solve a sequential decision problem such as finding the shortest route to reach a specific destination [35].

DP can deal with discrete variables, non-convex, non-continuous and non-differentiable functions. It is considered as a powerful optimization tool due to its ability to take into account the stochastic variability of the objective function. However, high computational requirements could reduce its efficiency in optimization problems requiring solutions in real-time. These computational problems are mainly related to the storage of information, especially when there is a large set of decision variables involved in the optimization problem.

Simulated Annealing

Simulated Annealing (SA) is a type of global heuristic search involving some random element in the way the algorithm proceeds [36]. SA is based on an analogy with a highly successful Monte Carlo simulation model for the physical annealing process of finding low energy states of a solid. Physical annealing is the process of finding the ground state of a solid, which corresponds to the minimum energy configuration by initially melting the substance and then lowering the temperature slowly, spending long times at temperatures close to the freezing point. In simulations of physical annealing processes, the system is given a small temperature displacement in any iteration before the resulting change in system energy is calculated. If the change results in lower energy, it is accepted. If not, it is accepted with a probability that is a function of the energy change and the temperature.

By analogy, the different solutions of an optimization problem can be seen as corresponding to the different states of a physical substance, and the cost function to be minimized as the energy of the system. Since the concept of temperature has no obvious equivalent in optimization problems, it is converted into a control parameter having the same units as the optimization function. In such a case, the search process will start at a given high "temperature" and the temperature is lowered every iteration. At each temperature, a certain number of changes will be considered. An acceptance function that defines the probability of accepting a change in system status that result in a higher objective function value will be used to accept or reject the change. Acceptance functions are generally defined so that the probability of accepting a change triggering an increase in objective function will be high at high temperatures but close to none when the temperature approaches zero.

One of the main advantages of the simulated annealing approach is that it avoids getting trapped at a local optimum by sometimes accepting a neighborhood move that increases the value of the objective function. However, one of the disadvantages is that it does not work in parallel, which means that the algorithm should finish the first stage completely before going to a second one.

Genetic Algorithm

A genetic algorithm (GA) is a heuristic near optimal solution search method based on natural genetics and mechanisms of natural selection [37]. Genetic algorithms are a particular class of evolutionary algorithms that use techniques inspired by evolutionary biology such as inheritance, mutation, selection, and crossover. As in biological evolution, organisms in a certain population have different sets of genes. Each gene corresponds to a decision variable of the optimization problem. As such, a set of genes held by an organism can thus be seen as representing a potential solution to the optimization problem at hand.

GA determines how well an individual's decision variables solve the problem at hand by evaluating a cost function (or objective function). Individuals with lower cost values (or higher fitness values) are seen as representing better solutions. The idea behind the genetic algorithm approach is to use an existing population of organisms to iteratively generate new organisms that may exhibit lower cost (or higher fitness). This improvement process is executed through the application of a series of genetic operators creating a new population of organisms on the basis of an existing population. The goal of this evolutionary approach is to produce diversity within the population and to explore various gene combinations or solutions to the optimization problem. However, this is not a blind search as the probability of reproduction of a specific organism is related to its cost function value. In other words, better parents are generally made to produce more children, thus theoretically increasing the chances of creating new organisms with lower cost function values.

Typical genetic operators used in GA include selection, crossover and mutation. While other operations may also be used, the three mentioned are the most commonly used. The functions performed by these operators are described in more details below.

Selection

Selection is a process in which individual strings (chromosomes) are copied into the new population according to their cost function (or objective function) value. Strings with a lower cost function value have higher probability of contributing one or more offspring in the next generation. The selection process is probabilistic. The probability for selection is based on the chromosome normalized cost function value relative to the normalized cost function values of all of individuals in the population. Selection begins by determining a chromosome's relative cost function value. This is done by dividing its cost function value by the sum of all the cost function values for all the chromosomes in the population. Then, a random number generator is used to select chromosomes for the crossover phase. The probability for a chromosome to be selected during each spin of the random number generator is equal to the chromosome's relative cost function value. The number of chromosomes selected is equal to population size, therefore, keeping the size constant for every generation. Some chromosomes may also be selected more than once, resulting in multiple copies of this chromosome to be used by the crossover operation [38].

Crossover

Crossover is the process of combining the genes of one chromosome with those of another to create offspring inheriting traits of both parents. The crossover rate is the likelihood of a chromosome being selected for the crossover operation. The chromosomes that are not selected will not have their genes changed before proceeding to the mutation phase. Those that are chosen will be paired with another chromosome that was also selected for crossover (a mate). From each pair, two offspring will be created to replace their parents. For classical one-point crossover, to determine which genes are inherited from parents, a random number between one and the total number of genes minus one is generated. For the first offspring, the genes numbered between one and the random number will be inherited from the first chromosome, and the genes numbered between the random number plus one and the maximum number of genes will come from the second chromosome. The genes for the second offspring will be inherited just like those of the first offspring except that the genes that came from the first chromosome in the first offspring will come from the second chromosome and those inherited from the second chromosome will come from the first chromosome [38].

Mutation

Just as in nature, some chromosomes will have random mutations occur in their genes. The mutation rate specifies the odds that a given gene in a chromosome will be mutated. If a gene is selected for mutation then its value will be changed. In the case of binary representation, the gene will simply be flipped (i.e., gene with a value of 1 changed to 0, or a gene with a value of 0 changed to 1 [38].

4.2.2. Optimization Method Selection

In this research, genetic algorithm optimization technique is used to optimize the cost function. The rationale behind this selection is due to the flexibility of representing a cost function using Genetic Algorithms and the ability of GAs to optimize complex functions. The GA allows more flexibility in formulating extremely complicated cost functions than other optimization methods. This flexibility includes ability to represent cost functions with discrete or continuous variables. Since the GA is only interested in the cost function value for a string solution, its performance is superior in avoiding getting trapped in local minima. Furthermore, the GA has the ability to deal with large computational problems since it can be run in parallel, and hence, the required computational time can be significantly reduced. For the problem in hand, if three cycles were optimized, there are 5.4566×10^{67} possible solutions, which required a long amount of time to find the optimal solution, where the GA can provide a near-optimal solution within an acceptable period of time.

4.2.3. Transit Dwelling Activity Prediction Model

This research requires developing a prediction tool for providing reasonably reliable information about the time elapsed between the moment a bus is first enters the bus stop and the moment it is ready to leave the bus stop, where dwell time activities may take place and result in variation around its average value. This section describes the possible approaches that can be adopted to develop the prediction tool and the rationale behind the selection of an ANN prediction method for the purpose of this research.

Regression Analysis

Regression Analysis is a statistical approach of modeling the relationships among three or more variables. It is used to predict the value of one variable (known as the criterion or dependent variable) given the values of others variables (known as predictors or independent variables). In this case, it is believed that variations in the values of the predictor variables cause or agree with variations in the values of the criterion variable. The objective of the regression is to evaluate the coefficients of an equation that is used to predict values of the criterion variables based on a specific combination of one or more predictor variables. The regression analysis assumes that the error, which is the difference between the predicted and measured values of the criterion variables are independent or each other, have a zero mean, have a constant variance across all values of the predictor variables, and are normally distributed. The violation of any of those assumptions leads to incorrect model structure [39].

In most cases, the equation that relates criterion and predictor variable to each other should be formulated first. This potentially limit the usefulness of this prediction method as results will be dependent on a priori choices regarding the set of predictor variables to consider [39]. This may create significant difficulties in problem considering large number of variables in which not only linear combinations are possible.

Artificial Neural Networks

Artificial Neural Networks, (ANNs) are non-linear statistical data modeling tools that are used to model complex relationships between input and output data or to find patterns in datasets. Instead of following a set of instructions to solve a problem or determine an outcome, ANNs take a non-algorithmic approach to problem solving by attempting to process information in a way similar to how the human brain operates. Figure 4.1 illustrates a typical ANN framework for a multilayer perceptron model [40]. ANNs have been used in many transportation-related applications, such as travel time prediction, crash rate modeling, and pavement distress prediction [41-43]. As illustrated, ANNs typically consist of three layers. Similar to the human brain, each layer is composed of a large number of highly interconnected processing elements called neurons. These neurons work in parallel to solve a specific problem or make a prediction [40, 44].

Within the framework of Figure 4.1, an outcome is generated from a series of input patterns $X_1 \ldots X_i$ subjected to transformation using functional information contained within the neurons and the communication channels (links). For the ANN supervised learning process, the target, or correct output has to be known. The goal of this type of network is then to correctly map the input $X_1 \ldots X_i$ to the corresponding output. The training takes place through repeated applications of all training input patterns and adjusting the strengths of the connections until the network's outputs approximately matches the correct or target outputs. In that sense, ANNs learns from examples, as children learn to distinguish dogs from cats based on examples of dogs and cats.

Since ANNs learn by example, the training examples must be selected carefully. Proper training is important, as well-trained ANNs may exhibit some capability for generalization beyond the training data, such as producing approximately correct results for new cases that were not used for training.

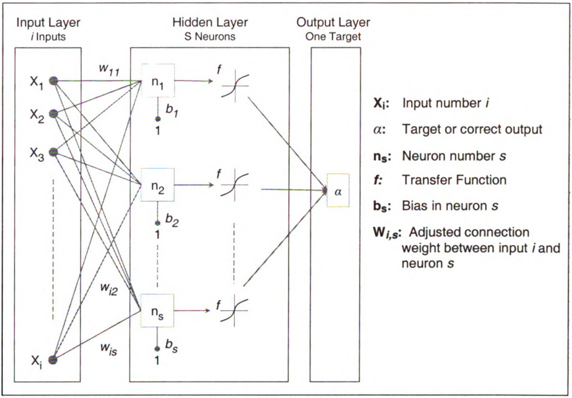


Figure 4.1 Typical architecture of a one-hidden-layer ANN with *i* input elements and *s* neurons follow.

<u>**Prediction Method Selection**</u>

For this research, the ANN model will be used as the basis for predicting the transit time elapsed between entering the bus stop and the time it is ready to leave the bus stop. This ANN prediction model is integrated within the D-SPORT controller. The selection of the ANN to perform the prediction can be explained by the fact that, unlike univariate and multivariate regression models, ANN are particularly well suited for approximating complex functional relationships, can handle noisy data, do not require predefined formulas or equations, and can handle intercorrelated predictor variables.

4.2.4. Simulation Environment

Use of computer simulation models is very common in the transportation field to analyze, assess, and evaluate different transportation systems and their components. In the world of traffic, simulation can be defined as a numerical technique for the conduct of experiments on digital computers. These simulation models may include stochastic characteristics, be microscopic, mesoscopic, or macroscopic in nature, and involve mathematical models that describe the behavior of a transportation system over extended periods of real-time [45]. The nature of this research requires evaluation of the operational effectiveness of traffic signal control algorithms in a variety of traffic situations. One approach would have been to evaluate the effectiveness of the proposed model through a field test. However, such an approach would not only require the purchase and installation of a variety of necessary and costly equipment, but may also limit the number of scenarios that could possibly be considered. Because of this, a need was identified for using a microscopic simulation environment to help develop and evaluate the proposed TSP algorithms. In this case, use of a simulation would allow one to quickly test the effectiveness of proposed algorithms, in addition to allow a wide variety of traffic conditions to be considered.

There currently exist a variety of microscopic traffic simulation packages used by traffic engineers to evaluate traffic signal control systems. Among the most commonly used are CORSIM [46], AIMSUN/2 [47], PARAMICS [48], and VISSIM [49]. Each simulation package has its own advantages and disadvantages. For the current research, a model allowing its user to define custom traffic signal control functions is needed. This immediately removes CORSIM from the list as this model does not allow its user to specify custom controller logic. AIMSUN/2 allows the user to develop a built-in traffic controller within the simulation package itself, however, the user is unable to attach an external controller within the simulation environment. PARAMICS and VISSIM are the only known simulation packages that allows the user to use an external controller developed using specific programming languages such as C, C++, Visual Basic and FORTRAN 90/95 [50]. In both cases, the user can use either the compiler provided with the simulation packages or Dynamic-Link Library (DLL), which is a collection of executable functions or data that could be shared between applications and are executed when needed.

A literature review has shown that VISSIM has been widely used to assess and evaluate traffic controllers implementing TSP algorithms [19, 51]. Based on this fact and the model's modeling capability, VISSIM was therefore chosen as the model that will be used in this research to assist with the development and evaluation of the proposed TSP model. Selection of this software is based more particularly on its ability to provide detailed modeling of transit operations and priority systems, including:

- Ability to model transit dwell times as a normal distribution with customizable parameters, a user-defined empirical distribution, or a function of the number of passengers waiting to board and alight at a particular stop.
- Provision of transit signal priority functions similar to those found in the real-world with the use of the National Electrical Manufacturers Association (NEMA) controller modeling.
- Ability to develop custom vehicle-actuated traffic signal control logic.
- Possibility to assign detailed priority rules at intersections and around bus stops.
- Ability to view the bus detections and resulting signal timing alterations.
- Ability to implement external controller within the simulation using the DLL.

4.2.5. Overall Research Approach

The proposed controller will use genetic algorithms to search for a near-optimal signal timing solution considering prevailing traffic conditions and fluctuations. As schematized in Figure 4.2, VISSIM will be set up to provide the traffic signal controller with current traffic and transit data in a real-time manner. This information will also be passed to the ANN prediction model, which will use it to predict transit dwell times and provide the result of this estimation to the controller. The controller will then use all available data to determine near optimum phase splits, cycle length, and offsets with adjacent intersections satisfying as best as possible both the transit and general traffic needs. Results of the optimizations could then finally be passed to the simulation model to test their impacts on transit and traffic operations.

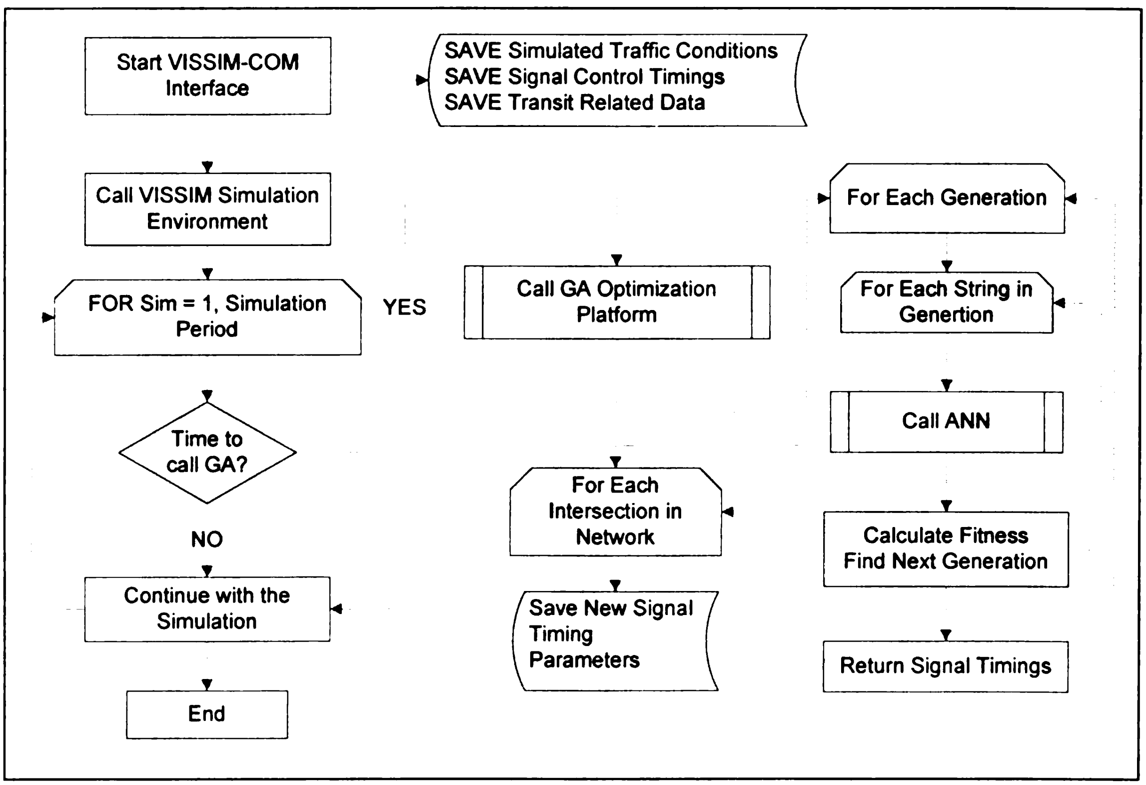


Figure 4.2 Schematic sketch for the D-SPORT controller element integration

To achieve the research objectives, the following sequence of research activities was proposed:

- Development of an artificial test network in the VISSIM microscopic traffic simulation environment is used for evaluating the D-SPORT controller. VISSIM is responsible for performing the traffic simulation, collecting and summarizing the simulation results, and making these results available to the D-SPORT controller upon request.
- Development of the D-SPORT controller cost function that addresses the issues discussed in Chapter 3. This work involves in delay and number of stops modeling for transit and non-transit vehicles travel throughout the traffic network, transit travel time modeling to predicting transit vehicle locations in the near future transit arrival time at bus stops.
- The use of an artificial neural network modeling approach to develop a tool for predicting bus arrival times at signalized intersections that can consider situations in which bus arrival times may be subject to significant variability due to either bus stop location or variability in bus dwell times.
- The use of a genetic algorithm optimization model to optimize the D-SPORT controller cost function (which includes the ANN model) and to return the near-optimal signal timings that can be applied at the signalized intersections operated in coordinated networks.
- Integration of the D-SPORT controller within the VISSIM microscopic traffic simulation model to be used as a research simulation environment using the

VISSIM-COM interface to allow the D-SPORT to communicate with the simulation environment.

- Utilization of the microscopic simulation model to validate the transit travel time prediction model proposed under the D-SPORT controller cost function, and to quantify the potential benefits of the D-SPORT controller with respect to providing TSP using current state-of-the-art techniques.

CHAPTER V.

DWELL TIME IMPACT ON TSP OPERATIONS

This chapter reports on a simulation study that was done to investigate the performance of a signal control system offering priority to buses approaching an intersection with a near-side stop under various traffic conditions. The objective of this work was to examine the potential impacts of dwell time variability on priority system performance and assess the importance of these impacts for the development of effective TSP control.

5.1. Simulation Modeling

This section describes the VISSIM simulation model that was developed to evaluate the impacts of dwell time variability on the performance of traditional active TSP systems offering red truncation (RT) or early green recall, and green extensions (GE) based on detection of approaching transit vehicles. The following subsections detail more specifically the geometry of the test network, the traffic signal control assumptions, the modeling of traffic and transit demands, the simulated TSP logic, the evaluation setup and the performance measures considered.

Test network

The test network is illustrated in Figure 5.1. This network includes two intersections located 305 m (1000 ft) apart and linked together by an arterial segment featuring two lanes per direction, ideal lane widths, no grade, no curb parking, left-turn bays, and a speed limit of 65 km/h (40 mph). The cross-streets feature similar characteristics, except for shared through-left lanes. Within this network, priority is to be provided exclusively to buses reaching Intersection 1 from the west. Intersection 2 is modeled to generate realistic platooned arrivals on the eastbound (EB) approach to Intersection 1.

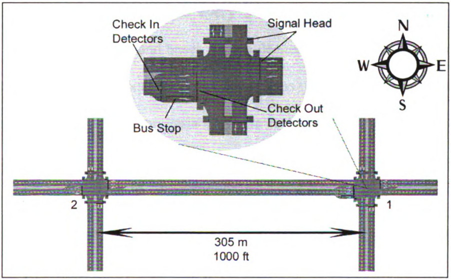


Figure 5.1 Simulated traffic network geometric layout

Traffic Demand

To remove truck effects, only passenger cars are simulated. Vehicles are generated at the upstream end of each entering link following Poisson distribution and VISSIM's default vehicle characteristics. Depending on the scenario, traffic volumes are set to produce equal volume-to-capacity (v/c) ratios of 0.50, 0.70, 0.80, 0.85, 0.90, and 0.95 on both the arterial and cross-street approaches. In each case, the appropriate volumes to assign to each approach were determined using the SYNCHRO 6 signal optimization software.

On arterial approaches with exclusive turning lanes, right and left turn flows are further assigned so as to maintain a v/c ratio of 0.50 on each turning lane. For the cross-streets, the left-turn and right-turn flows were simply assumed to correspond to 5% and 10% respectively of the total approach volume.

Transit Service

Buses are assumed to travel only in the eastbound (EB) direction and serve a single stop located in the right-turn bay of the eastbound approach to Intersection 1 with a normally distributed frequency of one vehicle every 5-min on average, and 1-min standard deviation. Both general traffic and buses share a desired 65 km/h (40 mph) travel speed. Transit vehicles entrance frequency, the interactions with other vehicles and traffic signals, in addition to dwell time variability, create fluctuations in the time individual buses take to reach Intersection 2. Since not all buses arrive at the same moment within the signal cycle, not all vehicles require a priority treatment. In the evaluated scenarios, 12 to 20 of the 23 buses generated over a 2-hour simulation end up requesting priority actions.

For the study, it was determined that a bus stop located at the intersection stop line would provide the most difficult priority case as such a stop would provide virtually no advance warning of the readiness of a bus to enter an intersection once boarding is completed. Starting from this premise, a stop located on an exclusive right-turn lane was then selected to minimize the impacts of dwelling buses on through traffic. Such a placement requires buses to wait for any queued vehicles on the right-turn lane to be cleared before being able to reach the stop. A similar situation would have existed were buses required to stop on a shared through/right-turn lane, expect that buses would in this case be waiting within the queue and not on an adjacent lane. To minimize the impacts of gap-seeking delays when leaving the bus stop, all buses are finally assumed to have priority of movement over the adjacent traffic, which is an assumption that can be regulated in the real world. While this facilitates the bus stop departure process, it does not completely eliminate the need to seek gaps in the adjacent lane. For instance, buses still have to wait for queues of vehicles on the adjacent lane to dissipate.

Scenarios featuring either stochastic or deterministic dwell times are considered. In the deterministic scenarios, every bus is assumed to dwell for exactly *n* seconds. In the stochastic scenarios, dwell time is instead a function of the number of individuals waiting to board a bus at the time of its arrival. Here, boarding demand is modeled by simulating the process of passenger arrivals at the bus stop. Depending on the scenario, a constant arrival rate ranging between 25 and 150 passenger/hr is considered, with an increment of 25 passenger/hr. Individual bus dwell times are then determined by assuming that 3-s are required for each passenger waiting to board, with another 3-s for door opening and closing. Bus occupancies were randomly assigned to each bus individually, and passengers are alighting in a rate of 10% of bus occupancy. Dwell time variations thus result from randomness in passenger arrivals at the bus stop combined with the effects of early/late bus arrivals.

Additional delays due to the need to wait for queued vehicles to clear the right turn lane and to wait for acceptable gaps in the adjacent traffic are not included in the dwell time estimates as such delays are more closely associated with traffic interactions than dwell activities. Ideally, however, such delays should be accounted for in algorithms used to predict transit travel time throughout the traffic network.

Traffic Signal Control

All traffic signals are controlled on a fixed-time basis. Each intersection is controlled with a 90-s cycle and two phases implementing a 50:50 green split between the east-west and north-south travel directions. For each phase, clearance intervals are modeled by 4-s amber followed by a 1-s all-red. A 20-s minimum green is further imposed to ensure that minimum pedestrian crossing times are not jeopardized by priority actions. Signal coordination between the two intersections is obtained from optimizing each experiment using SYNCHRO 6, a software package for modeling and optimizing traffic signal timings [52] by taking into account the traffic volume and associated with each experiment.

Priority Setup

Transit movements are monitored through two detectors. A first detector is placed upstream of the stop line to detect approaching buses when they are 3-s away from the stop line in situations in which their progression is not impeded by other traffic or a need to serve boarding/alighting passengers. The second detector is placed just downstream of the stop line to determine when a bus has entered the intersection and no longer requires priority.

Priority actions are provided using Vehicle Actuated Programming, (VAP), which is a VISSIM's programming module. This module allows the building of custom controller functions based on information provided by modeled traffic detectors. As indicated by the following rules, the simulated TSP logic is designed to provide green extensions and early recalls based on the expected stop line arrival time of detected buses, a maximum allowed green extension, and minimum green constraints:

- For the fixed dwell time scenarios, the signal controller predicts stop line bus arrivals using a fixed dwell duration corresponding to the actual vehicle dwell time.
- For the stochastic dwell time scenarios, the signal controller is not assumed to know the fixed dwell time of each bus. Priority needs are assessed by simply considering an average dwell duration based on the average number of boarding passenger expected at the bus stop under a perfectly regular service.
- No priority action is taken for projected arrivals during the normal green.
- For projected arrivals during the red, green extensions are granted only if buses are arriving close to the end of the green, within the maximum allowed extension window.
- Red truncations (RT), or early recalls are considered for projected arrivals during the red beyond the maximum green extension window, but only for arrivals close enough to the end of the red. For a maximum extension of 10 s, an early recall will for instance be considered only for arrivals within 10-s of the end of red.
- Green extensions (GE) and early recalls are granted in increments of 5-s, with a maximum of 5 to 20 seconds depending on the scenario.

- Any granted extension is terminated if a bus is detected to have crossed the stop line.
- Each bus is allowed only one priority request. A bus failing to cross the stop line in time to benefit from a green extension will thus not be considered for an early recall.

Evaluation Setup

Evaluations are based on two-hour simulations. For each simulation, the first 5 min is neglected to allow network loading. Past this interval, traffic data are collected in five-minute intervals, while transit data are collected for individual buses. To account for the stochastic nature of VISSIM's simulation processes, 10 replications are made of each scenario using different random number seeds. Unless otherwise noted, all results reported are thus for an average of 10 replications.

Figure 5.2 summarizes all the possible experiment combinations that are examined under each priority setup. The total of 360 scenarios are tested, 180 are assigned to the fixed dwell time priority setup, and 180 for priority setup with variable dwell time. The differences between these scenarios are by combining volume-to-capacity ratios along the major street (i.e., the Eastbound (EB) and Westbound (WB) directions), with maximum allowable signal alterations for green extensions and early green recalls, and the average passenger demands at the bus stop.

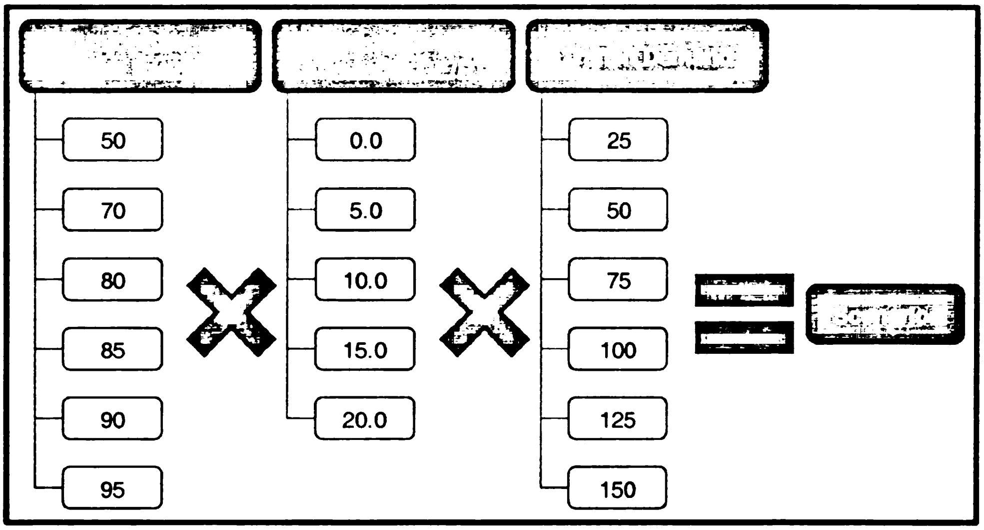


Figure 5.2 Sketch summarizes the scenarios that are examined under fixed and variable dwell time priority setup

Performance Measures

Two performance measures were retained to assess the impacts of dwell time variability on transit and traffic operations. The first is the average delay per vehicle. This delay is calculated by VISSIM as the difference between the actual travel time experienced by a vehicle and the travel time that would have been observed if the vehicle had been able to travel at its desired speed without any impedance. The same definition is used for buses, except that dwell time is excluded from the calculations as this is considered a normal transit activity. Delays associated with buses slowing down on the approach to a stop and bus stop clearance times are considered as travel time delays.

The second measure is priority system efficiency. This parameter measures the ability of buses requesting priority to benefit from the system based on the following criteria:

- Buses arriving during the normal green are not considered as they do not require priority.

- Buses requesting priority and entering the intersection during a provided green extension (GE) or early recall (RT) are counted as successes.
- A bus requesting a green extension and arriving too late to benefit from it is considered as a failure.
- A bus requesting a green extension but arriving earlier than expected and for which a priority request cancellation can be issued before the extension is implemented is considered as a success since the system was able to adjust to the early arrival.
- The awarding of early green recalls for buses that end up arriving during the normal green is considered as a success as reduced delays may still result from early queue dissipation.

At the end of a simulation, the efficiency is determined by the ratio of the number of buses having benefited from the system over the total number of buses having requested a priority action.

5.2. Impacts of Dwell Times on Transit Operations

Table 5.1 compares the average bus delays obtained under deterministic and stochastic dwell times for various combinations of v/c ratio, passenger boarding demand, and maximum limits on green extensions/early recalls. For each combination of v/c ratio and passenger boarding rate, it is first observed that allowing longer signal alterations tends to reduce bus average delays for both the deterministic and stochastic dwell time assumptions. Transit delay reductions reach up to 40 % when considering a 20-s limit on signal alterations with respect to no priority action. This was expected as longer extensions provide more flexibility for accommodating buses. For a given alteration limit, v/c ratio, and dwell time, it is further observed that increases in the number of boarding passengers tend to decrease average delays. This is an effect directly associated to the geometry of the test network and the traffic signal setup for the scenarios considered. For instance, longer dwell times could push more buses to reach the prioritized intersection during the normal green, when no delay is incurred.

The results of Table 5.1 indicate that dwell time variability clearly impacts the ability of buses to benefit from a priority system. As observed, stochastic dwell scenarios produced lower average bus delays than deterministic scenarios. This is contrary to what was expected. These delay reductions can be explained by distributing bus arrivals at stop line under variable dwell times, where the probability of successfully using a granted TSP treatment.

To illustrate this point, consider the sketch of Figure 5.3. This sketch illustrates two scenarios of potential bus arrival times. The first scenario illustrates a situation in which buses arrive in the middle of the red signal after having dwelled for a certain time upstream of the intersection. In the absence of variability, all buses would reach the intersection at exactly the same moment, as exemplified here by the thick trajectory line. This means that all buses would receive the same priority treatment and incur identical delays. However, as dwell time variability spreads arrivals over a wider window, some buses would start to arrive close enough to the red initiation to be able to benefit from a green extension. Buses arriving closer to the end of the red would also incur less delay. At the system level, these reduced delays due to increased variability of arrival would then translate into lower average bus delays for the intersection.

Table 5.1 Average transit delay for the scenarios considered, in s/veh

Passenger / Hour	Fixed Dwell Time (s) Maximum GE/RT Alteration					Stochastic Dwell Time (s) Maximum GE/RT Alteration					Difference (%) Maximum GE/RT Alteration				
	0	5	10	15	20	0	5	10	15	20	0	5	10	15	20
v/c: 0.50															
25	38.0	35.4	31.6	29.2	25.8	33.2	31.3	27.7	25.3	22.5	-12.6	-11.6	-12.5	-13.6	-12.5
50	32.7	30.2	26.5	23.4	20.5	29.2	26.6	23.3	20.5	18.1	-10.8	-11.9	-12.3	-12.7	-12.0
75	28.6	25.9	22.5	20.1	17.1	26.6	24.0	20.3	18.0	15.9	-6.8	-7.5	-10.0	-10.4	-7.5
100	26.3	21.9	19.5	17.0	15.0	24.3	20.5	18.4	16.5	15.3	-7.5	-6.7	-5.4	-3.1	1.5
125	22.7	20.3	17.7	15.7	15.0	21.6	19.0	17.1	16.0	15.5	-4.8	-6.5	-3.8	1.7	3.3
150	19.7	17.3	15.6	15.4	15.4	20.1	18.3	16.7	16.0	15.9	2.1	5.6	6.9	4.2	3.4
v/c: 0.70															
25	40.3	36.7	33.5	30.3	27.1	36.5	32.4	29.2	26.0	23.5	-9.6	-11.6	-12.6	-14.2	-13.4
50	35.4	32.1	28.9	25.5	22.9	33.0	29.5	26.0	22.7	20.9	-6.9	-8.2	-9.9	-10.7	-8.7
75	32.2	29.2	25.2	22.7	19.9	32.0	26.5	23.3	21.4	19.4	-0.7	-9.1	-7.7	-5.6	-2.3
100	32.1	26.5	24.8	23.0	20.9	29.2	25.0	23.2	21.2	19.7	-9.2	-5.7	-6.6	-7.6	-5.8
125	29.5	26.0	24.8	23.0	20.9	26.9	24.5	22.6	23.3	21.3	-8.6	-6.1	-8.9	1.3	1.7
150	25.7	24.7	21.9	21.7	20.5	24.5	22.5	22.6	20.7	21.0	-4.8	-8.9	3.1	-4.5	2.1
v/c: 0.80															
25	41.0	38.1	35.7	32.7	30.7	36.7	35.2	31.6	29.6	26.8	-10.7	-7.6	-11.3	-9.4	-12.7
50	36.5	33.7	30.5	28.4	24.9	35.4	32.3	29.6	25.7	24.1	-3.0	-4.2	-2.9	-9.4	-3.4
75	35.3	30.6	28.0	26.3	23.6	34.8	32.1	26.9	24.8	21.2	-1.6	4.9	-3.8	-5.8	-10.4
100	35.5	32.8	29.0	25.9	23.8	33.6	32.0	27.6	24.7	23.1	-5.2	-2.3	-4.9	-4.7	-2.9
125	34.4	31.9	30.6	27.9	27.9	32.5	30.3	28.1	26.2	24.6	-5.6	-5.0	-8.2	-6.3	-11.6
150	30.9	28.6	27.4	25.9	25.2	29.0	26.8	28.4	25.0	24.2	-6.3	-6.3	3.8	-3.3	-3.7
v/c: 0.85															
25	42.1	39.1	36.3	33.5	30.3	40.3	35.8	32.4	30.0	27.2	-4.3	-8.5	-10.7	-10.3	-10.2
50	37.8	35.2	32.6	30.0	26.8	38.5	34.7	30.1	26.8	24.8	2.1	-1.3	-7.7	-10.8	-7.4
75	37.2	34.3	30.6	27.5	26.7	38.9	33.7	30.0	27.3	23.7	4.5	-1.7	-1.8	-0.7	-11.2
100	40.0	36.0	32.3	28.9	28.0	36.6	34.6	31.4	28.1	26.6	-8.4	-3.7	-2.7	-2.8	-5.1
125	37.4	36.3	33.0	30.8	30.2	35.1	33.0	30.5	28.4	28.4	-6.3	-9.1	-7.5	-7.6	-6.0
150	32.5	31.4	29.4	28.1	27.6	32.7	31.3	28.6	28.8	26.2	0.6	-0.3	-2.8	2.4	-4.9
v/c: 0.90															
25	43.1	40.1	37.0	34.2	31.3	41.4	37.0	33.1	31.6	29.4	-3.8	-7.9	-10.5	-7.5	-6.0
50	39.5	36.1	34.1	30.6	28.5	42.0	35.4	31.3	28.5	26.0	6.3	-1.8	-8.3	-6.9	-8.7
75	40.1	34.8	31.3	28.6	26.8	40.1	34.8	31.9	28.7	25.0	0.1	0.1	2.2	0.3	-6.5
100	44.3	35.4	32.8	30.7	28.3	38.9	35.4	31.6	28.5	27.3	-12.2	0.0	-3.4	-7.0	-3.2
125	38.8	37.3	32.2	30.8	28.9	33.9	33.6	31.4	29.9	28.6	-12.6	-10.0	-2.5	-3.1	-1.0
150	37.3	33.9	31.3	30.5	28.6	32.8	31.8	30.6	29.1	27.7	-12.1	-6.0	-2.4	-4.4	-3.1
v/c: 0.95															
25	43.1	40.8	37.9	35.3	32.2	42.4	37.2	34.7	32.2	30.0	-1.8	-8.8	-8.3	-8.8	-7.1
50	41.6	37.4	34.4	30.6	29.2	42.3	37.7	32.0	29.6	27.5	1.8	0.8	-7.2	-3.5	-5.9
75	40.6	35.3	32.9	30.5	27.7	41.5	37.0	32.7	29.4	28.2	2.3	5.0	-0.5	-3.4	2.0
100	43.2	38.5	37.9	32.5	30.7	40.8	37.1	34.5	30.7	29.5	-5.4	-3.8	-9.0	-5.4	-3.8
125	40.1	38.9	35.2	35.6	31.4	37.2	36.8	31.6	29.8	29.0	-7.2	-5.3	-10.5	-16.3	-7.9
150	36.2	33.6	32.2	31.3	29.8	34.0	34.3	31.6	30.2	28.6	-6.2	2.2	-1.8	-3.5	-3.7

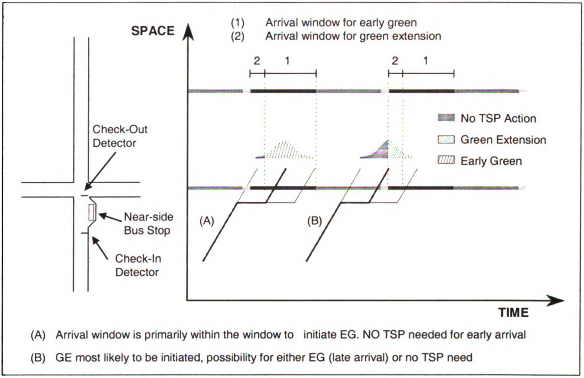


Figure 5.3 Likelihood of priority actions with respect to arrival window

Another effect is linked to the number of buses stopping to board passengers. Under fixed dwell times, all buses are assumed to service the simulated bus stop. However, under stochastic dwell times buses may skip the stop if no passenger is waiting to board. In such a case, buses skipping the stop would then tend to arrive during the normal green and incur no delay, thus reducing the estimated average incurred delays.

Figure 5.4 provides another look at the impacts of varying dwell times on bus delays. The figure illustrates the bus delays that were obtained for fixed and stochastic dwell times ranging virtually from 0 to 1.5 the cycle length. For each scenario, the standard deviation and average of all observed delays are shown. It is first noticed that stochastic dwell times tend to lead to higher average delays in some scenarios and lower delays in others. As explained earlier, these effects are due to variations in the proportion of arrivals during the normal green or at a moment for the application of priority treatment.

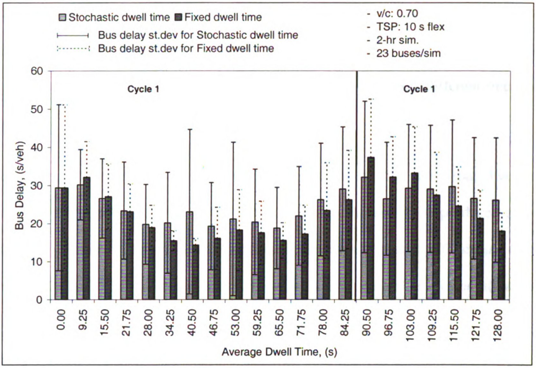


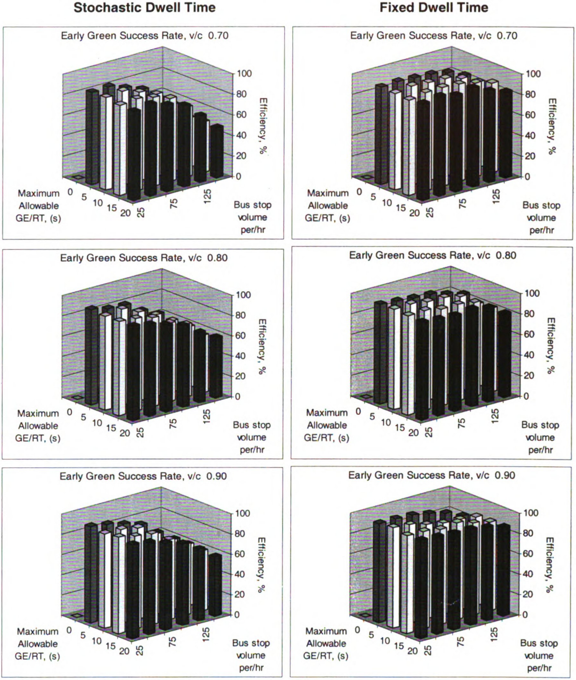
Figure 5.4 Average bus delay with respect to average dwell time and signal cycle length

A particular observation is that a cyclic pattern appears to exist between the average bus delay and the dwell duration with respect to the cycle length. Here, priority request issuance and incurred delays are at a minimum when bus arrivals generally correspond to the middle of the green signal. As bus arrivals shift towards the end of the green as a result of longer dwell times, an increasing proportion of buses start to arrive during the early portion of the red. While green extensions can initially accommodate early red arrivals, eventually more buses start to arrive past the point for which an extension can be provided and are thus forced to stop. This causes delay increases as the stopping buses must either wait for a time when an early green recall can be implemented or the entire red duration if a recall cannot be granted. Maximum delays occur when the average arrival time coincides with the beginning of the red signal. As the arrivals further shift towards the beginning of the green, an increasing proportion of buses start to arrive during the green or close enough to the end of the red to fully benefit from early green recalls, thus causing reductions in incurred delays. A perfectly symmetric impact cycle is not necessarily obtained as green extensions and early recalls have different impacts on delays.

With respect to system efficiency, Figure 5.5 and Figure 5.6 show that dwell time variability tends to reduce efficiency. While deterministic scenarios produced systems with 100% efficiency, corresponding stochastic scenarios produced efficiencies varying between 40% and 95%. This was expected as dwell time variability creates situations in which granted priority treatments may end up not being used by the buses that had initially requested them. When no bus has requested a green extension or early green request during the 2-hour simulation period, the system efficiency is reported as 0% since the system efficiency calculations are not applicable.

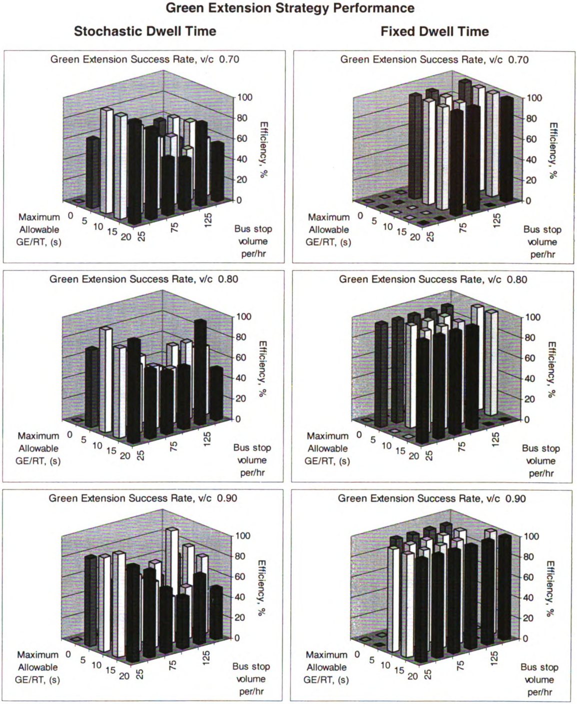
For a given boarding rate, system efficiency does not appear to be significantly affected by increasing the limit on timing alterations, especially when efficiency is already high. This is explained by the implementation of green extensions and recalls that are shorter than the maximum allowed for a number of buses. For these buses, increasing the alteration limit does not lead to changes in priority actions. For a given alteration limit, however, system efficiency tends to reduce with increasing boarding rates. This is again due to a change in the proportions of buses arriving during the green and red portions of the cycle.

Early Green Strategy Performance



* Zero values indicates that no bus was observed

Figure 5.5 Efficiency of early green strategy under different level of service



* Zero values indicates that no bus was observed

Figure 5.6 Efficiency of green extension strategy under different level of service

5.3. Impacts on Traffic along Priority Route

Figure 5.7 illustrates impacts on passenger car delays for vehicles traveling on the east-west arterial. It is observed here that dwell time variability has relatively small impacts on the general traffic. This is explained by the fact that dwell time variability does not translate into significant changes in priority actions in the scenarios considered. While variability leads to fluctuations in bus arrival times, priority needs are determined under both the deterministic and stochastic dwell time assumptions using similar expected average dwell time assumptions. In both cases, buses thus tend to request similar priority actions even though buses with stochastic dwell times will not necessarily all fully benefit from the granted green extensions and early recalls.

Figure 5.7 shows that increases in average boarding rate (i.e., increase in dwell time) lead to increased traffic delays, and that increases in the maximum signal alteration limit tend to reduce delays for a given dwell duration. In the later case, delay reductions are primarily due to the increasing ability of buses requesting priority to lengthen the effective green signal along the east-west arterial. In the former case, the delay increases with longer dwell times are again attributed to changing proportions of buses arriving during the green and red portions of the signal cycle. As the dwell time increases, more buses arrive during the green, thus reducing the need for priority and the opportunities for increasing the effective green along the arterial approaches.

Non-transit traffic tends to benefit the most with early recalls. When green extensions are awarded, vehicles are able to avoid delays corresponding to an entire red duration. However, only a few vehicles will benefit as the highest flows are generally observed at the beginning of a green phase. For an early green recall, the potential delay savings are limited to the duration of the recall, but all the vehicles in queue at the intersection at the time of the action will benefit from an earlier departure. Thus, despite yielding lower delays per vehicle, early recalls have a potentially greater impact on the overall traffic through their ability to reduce delays for a greater number of vehicles. However, it should be kept in mind that these results do not account for potential delay increases that may result at downstream intersections from changes in progression patterns as the evaluation only considered an isolated intersection.

5.4. Impacts on Cross-Street Traffic

Impacts on cross-street traffic were found to be generally opposite to those for vehicles traveling along the prioritized bus route. While priority actions increased effective green time and reduced delays along the east-west arterial, these actions reduced green durations and increased delays on cross-streets. Whether a priority system will cause overall traffic delays to increase or decrease will then depend on the proportions of vehicles benefiting and suffering from the priority actions.

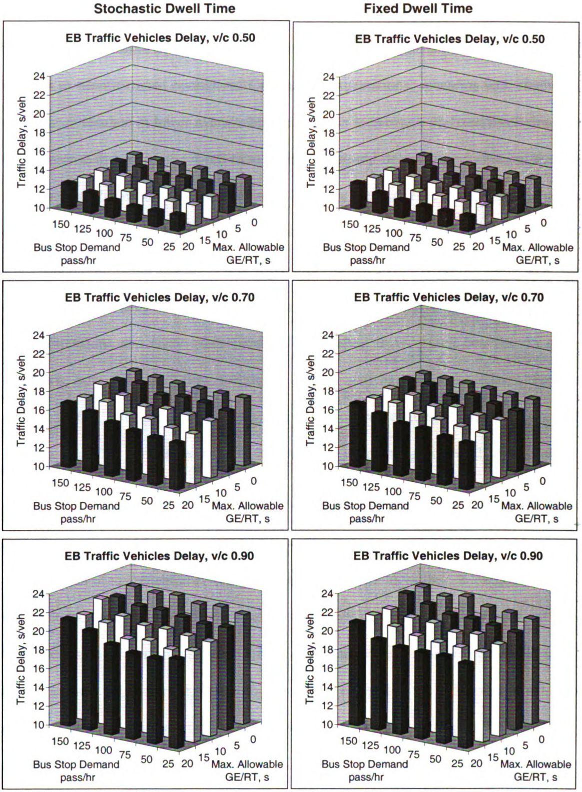


Figure 5.7 Traffic vehicle delays for scenarios considering stochastic dwell times

5.5. Summary

The simulation results indicate that dwell time variability impacts the ability of priority systems to benefit transit operations. Contrary to expectations, scenarios with stochastic dwell times resulted in lower average bus delays than scenarios with fixed dwell times. It was expected that increased dwell time variability would decrease performance. In both cases, no significant impacts were observed on the general traffic. The contrary impacts on transit operations are attributed to changes in bus arrival patterns within the simulated signal operation. In the simulated scenarios, dwell time variability pushes buses to arrive at moments when they can fully benefit from the ability of a signal system to implement green extensions and early green recalls, thus leading to increased benefits. In other situations, however, dwell time variability may as well have opposite effects. For the general traffic, the lack of noticeable impacts is the results of the fact that all implemented TSP actions were based on expected average transit dwell times and not actual dwell times.

Overall, the results of the simulation reported in this chapter indicate a need to retain consideration of the potential effects parameter in any TSP algorithm. In particular, it can be concluded that if the objective is to evaluate general traffic impacts associated with the implementation of a priority system, simulations or algorithms considering average dwell times can be used. However, if the objective is to determine the efficiency of selected priority strategies or algorithms with respect to transit operations, it then becomes important to consider dwell time variability in the TSP optimization.

CHAPTER VI.

D-SPORT COST FUNCTION FORMULATION

This chapter discusses in details the mathematical development of the cost function that will be fed into the GA optimizer to find a near-optimal timing plans with TSP actions. First, a brief description for the traffic signal control assumptions is introduced. Next, the formulations for the decision variables that will be used in the cost function are presented. The remaining sections show the mathematical derivatives for the general cost function and the sub-cost functions.

6.1. Traffic Signal Control Assumptions

To support the evaluation of the D-SPORT controller performance, a custom traffic signal controller will need to be developed for operation within the VISSIM microscopic traffic simulation model since the functions provided with the software do not provide the desired functionalities. As discussed earlier, the D-SPORT controller will have to provide TSP treatments for buses traveling in an urban coordinated signalized network with other traffic where some uncertainty may exist regarding true transit vehicle arrival times at the intersection stop line.

In developing the D-SPORT cost functions, the following assumptions are made:

- Transit vehicles are all equipped with a Global Positioning System allowing continuous tracking of their movements.
- All buses are equipped with an automatic passenger counter system.
- A continuous communication system exists allowing the signal controllers to obtain in real-time all relevant information about approaching buses (such as position, speed, and number of passengers).
- The D-SPORT controller is to operate in real-time.
- Minimum green times must be granted for all phases in order to accommodate pedestrians who are crossings the street.
- Maximum green times must not be exceeded for all phases
- The traffic signals operate with two phases only.

The traffic signals were set to operate with two phases only to simplify the problem at hand and to furnish the background required to develop a model that can eventually be improved to handle more complicated phasing schemes and traffic composition.

6.2. Decision Variables

The decisions variables upon which a signal controller operates are cycle lengths, green times, and offsets. Each of these variables is described in more details in the following paragraphs.

Cycle Length

The cycle length equals to the summation of all signal intervals within a complete signal display sequence. This includes the summation of green times, yellow and all-red intervals for all phases. Mathematically, this can be expressed by Equation 6.1 below:

$$CL_i = \sum_{n=1}^{\Phi}(g_{ni} + rr_{ni} + y_{ni}) \qquad \text{.................................. Equation 6.1}$$

Where:

CL_i:	Cycle length of intersection *i*, in seconds
g_{ni}:	Green duration for phase *n* at intersection *i*, in seconds
y_{ni}:	Yellow duration for phase *n* at intersection *i*, in seconds, set to 4 s
rr_{ni}:	All red duration for phase *n* at intersection *i*, in seconds, set to 1 s
Φ:	Total number of phase for intersection *i*

Green Time

For each traffic movement, the green time represents the interval during which traffic is allowed to enter an intersection. For the optimization at hand, this parameter is bound by both a maximum and a minimum value, as expressed in Equation 6.2:

$$g_{ni} \in [g_{\min}, g_{\max}] \qquad \text{.................................. Equation 6.2}$$

Where:

$g_{\min}$:	Minimum allowable green time, in seconds, set to 20-s
$g_{\max}$:	Maximum allowable green time, in seconds, set to 40-s

A minimum green time must be imposed to ensure that pedestrians are always given enough time to cross the intersection or to provide a guaranteed minimum traffic service time. On the other hand, a maximum green time must be specified to avoid allocating disproportionably long green intervals that may not be well received by motorists. To determine the lower and upper limits of a green duration, a careful consideration of the geometric layout of the intersection and the time required for pedestrians to cross the intersection and traffic conditions is required.

Offsets

An offset is defined as the time elapsed between the beginning of a green phase at an intersection, and the beginning of a corresponding green phase at the following intersection. An efficient offset would be the one allowing the first vehicle of a platoon leaving an intersection to reach the downstream intersection right when the tail of the downstream queue has just cleared. Such an offset would in theory maximize utilization of green time and contribute to reduced stops and delays.

To calculate the offset, information is required about the time taken by vehicles to travel from one intersection to the next. This travel time is function not only of the desired traffic speed but also of the initial speed at which vehicles enter the segment and the rate at which vehicles can accelerate and decelerate. Assuming a linear relationship between acceleration and speed, the following equation can be used to determine the maximum acceleration rate of a vehicle traveling at a given speed v [53].

$$acc = \upsilon + \tau \cdot v \qquad \text{................................ Equation 6.3}$$

Where:

acc : Vehicles acceleration rate, in ft/s^2

υ : Constant parameter representing the vehicle's maximum feasible acceleration, in ft/s^2

τ : Constant parameter representing the rate of reduction in acceleration with increasing speed, $\tau < 0$, in s^{-1}

v : Speed of vehicle, in ft/s

This equation correctly represents that a vehicle's acceleration capability tends to decrease with increases in speed. While this model is not perfect, it represents a commonly used assumption in the modeling of vehicle maximum acceleration capabilities [53]. In this case, the equation is used to model the actual acceleration of a vehicle and not necessarily the maximum acceleration capability. This is done by simply assuming that the actual vehicle acceleration can be expressed as a constant fraction of the maximum acceleration capability.

Equation 6.3 can be utilized to calculate the time required to travel across a link assuming no initial queue in the downstream intersection. Figure 6.1 illustrates the speed change over time for a vehicle traveling between intersections i and j. The vehicle leaves intersection i with an initial speed v_i at t_o, and accelerates until it reaches the desired speed v_f at t_1. After traveling at the desired speed, the vehicle reaches intersection j, at t_2. Therefore, the ideal offset between the two intersections should be synchronized with the arrival of the vehicle. The following equations show how the link travel time can be mathematically derived from Equation 6.3.

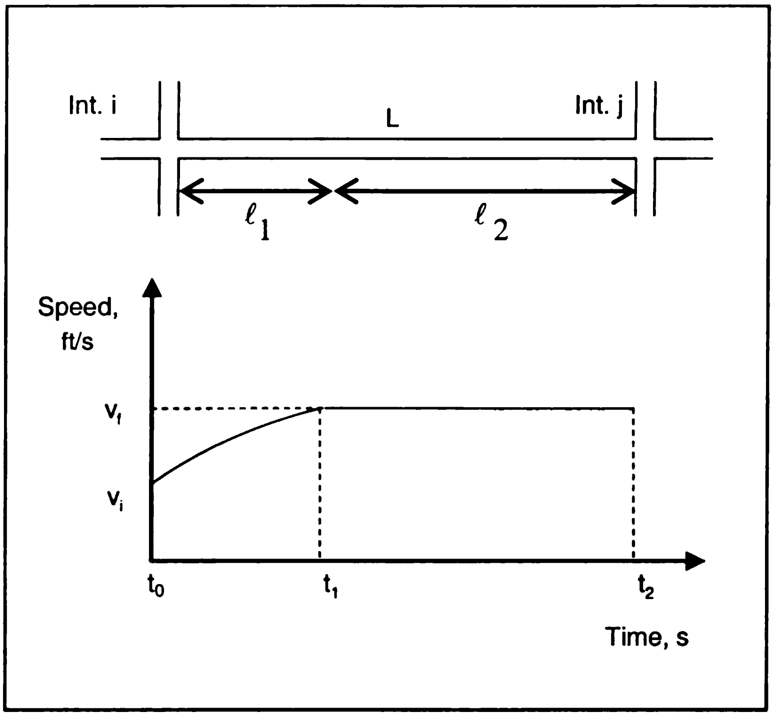


Figure 6.1 Schematic sketch for speed profile through the link

$$\frac{dv}{dt} = \upsilon + \tau v \quad \rightarrow \quad \int_0^t dt = \int_{v_i}^{v} \frac{dv}{\upsilon + \tau v} \quad \rightarrow \quad t_1 = \frac{1}{\tau} \ln\left(\frac{\upsilon + \tau v_f}{\upsilon + \tau v_i}\right)$$

$$t = \frac{1}{\tau} \ln(\upsilon + \tau v)\Big|_{v_i}^{v} \quad \rightarrow \quad t = \frac{1}{\tau} \ln\left(\frac{\upsilon + \tau v}{\upsilon + \tau v_i}\right) \quad \rightarrow \quad (\upsilon + \tau v_i)e^{\tau t} = \upsilon + \tau v$$

$$v = \frac{1}{\tau}\left[(\upsilon + \tau vi) \cdot e^{\tau t} - \upsilon\right] \quad \rightarrow \quad \int_0^{\ell_1} ds = \int_0^{t_1} \frac{1}{\tau}\left[(\upsilon + \tau vi) \cdot e^{\tau t} - \upsilon\right] \cdot dt$$

$$\ell_1 = \left[\frac{\upsilon + \tau vi}{\tau^2} e^{\tau t} - \upsilon t\right]\Bigg|_0^{t_1} \quad \rightarrow \quad \ell_1 = \frac{v_f + \tau v_i}{\tau} - \frac{\upsilon}{\tau} \ln\left(\frac{\upsilon + \tau v_f}{\upsilon + \tau v_i}\right)$$

$$But\ t_2 = \frac{L - \ell_1}{Vf} \quad \& \quad LinkTravelTime = t_1 + t_2 \quad \rightarrow$$

$$LinkTravelTime = \frac{1}{B}\ln\left(\frac{\upsilon+\tau v_f}{\upsilon+\tau v_i}\right) + \frac{L - \frac{v_f+\tau v_i}{\tau} - \frac{\upsilon}{\tau}\ln\left(\frac{\upsilon+\tau v_f}{\upsilon+\tau v_i}\right)}{Vf} \quad \Rightarrow$$

$$LinkTravelTime = \frac{L_{i\to j}}{V_f} + \left(\frac{\upsilon+v_i}{\tau\cdot v_f}\right)\cdot\ln\left(\frac{\upsilon+\tau\cdot v_f}{\upsilon+\tau\cdot v_i}\right) \quad \text{.....................} \quad \text{Equation 6.4}$$

Another parameter required to determine the ideal offset is the time required to clear a queue of vehicles that may be present at the downstream intersection. This parameter is important if it is desired to allow vehicles being released at an intersection not to be stopped by a dissipating queue while traveling to the next intersection.

Equation 6.4 calculates the ideal offset based on the time required to travel between intersections *i* and *j* under free flow conditions. However, this ideal offset should be adjusted to take into account any queue presented in the downstream intersection. The adjustment consists in subtracting from the ideal offset the time required to dissipate the queue before the upstream platoon arrives. The queue dissipation time is a function of number of vehicles in queue and the average discharge headway in saturated conditions. For a one way corridor with no allowed turning movements, the ideal offset can therefore be computed using Equation 6.5.

$$I\_off_{ij} = \frac{L_{ij}}{V_f} + \left(\frac{\upsilon+v_i}{\tau\cdot v_f}\right)\cdot\ln\left(\frac{\upsilon+\tau\cdot v_f}{\upsilon+\tau\cdot v_i}\right) - q_j\cdot\bar{h} \quad \text{.................} \quad \text{Equation 6.5}$$

Where:

I_off_{ij}: Ideal offset between intersections *i* and *j* for any given cycle, in seconds

v_i: The initial speed at which the first vehicle in the platoon enters the segment, in ft/s

v_f: The desired speed reached by the first vehicle in the platoon, in ft/s

q_j: Initial queue at intersection j during cycle c, in number of vehicles

$\bar{h}$: Average saturation headway, in s/veh, set to 2.5 s/veh

6.3. D-SPORT Controller Cost Function Formulation

The cost function can be broken down into a set of sub-functions each representing a specific objective. The combination of all these sub-cost functions into a single equation would then create the general cost function. Since different agencies may have different priorities and perspectives on what constitutes ideal TSP treatment, this approach would allow the optimization process to be tailored to the specific needs of each agency. For instance, if an agency is interested in implementing TSP in such a way that transit vehicles experience less delay, then the optimization process would be defined with a cost function minimizing delays. If another agency is interested in regulating transit schedule, the cost function considering this parameter specifically could instead be used within the same optimization framework.

To reflect the potential for considering alternate optimization objectives, the following cost function has been coded for optimization by the GA. This equation gives a weighted factor for each of the sub-cost functions. Attribution of a value of zero to any of the weighting parameters will effectively cancel consideration of this parameter in the optimization. Similarly, attribution of a value above zero will allow the parameter to be considered with higher or lower importance in the optimization depending on the values assigned to the individual weight parameters.

$$COST = \left[\sum PI \quad \sum TTT \quad \sum SCH\right] \cdot \begin{bmatrix} \alpha_{PI} \\ \alpha_{TTT} \\ \alpha_{SCH} \end{bmatrix} \quad \text{.....................} \quad \text{Equation 6.6}$$

Where:

$COST$:	Cost function value
PI :	Traffic network overall performance index, (defined in Section 6.3.1)
TTT :	Transit travel time (defined in Section 6.3.2)
SCH :	Schedule adherence penalty (defined in Section 6.3.3)
α_{PI} :	Traffic network performance index parameter sub-cost parameter, has a value of 1.0
α_{TTT} :	Transit travel time parameter sub-cost parameter, has a value of 10.0
α_{SCH} :	Schedule adherence sub-cost parameter, has a value of 10.0

In this research, the selection of the values for the sub-cost parameters (i.e., α_{PI}, α_{TTT}, and α_{SCH}) is based on the sensitivity of the cost function optimization to these parameters. Specifically, weight values were assigned to produce a cost function in which the magnitude of the performance index, travel time and schedule parameters are within the same scale.

The following paragraphs describe the mathematical derivation for each of the performance sub-cost considered in the general cost function.

6.3.1. Traffic Network Performance Index

The network performance index, (PI) is defined as the sum of the weighted linear combination of estimated delay and number of stops per vehicle per unit time for all signal phases under a given timing plan [54]. The PI will be used for optimizing the signal timing to accommodate non-transit vehicles. For an isolated intersection, ψ is the vector that expresses the signal timing parameters (i.e., the cycle length CL, green time g_{Φ}, phases vector Φ, and offset *off*). This can be expressed mathematically as:

$$\psi = [CL \quad g_{\Phi} \quad \Phi \quad off] \qquad \text{Equation 6.7}$$

For a given signal timings ψ, the PI can therefore be mathematically expressed as a function of:

$$PI = \sum_{\forall i \in INTX} \left(W_d \cdot D_i(\psi) + W_{ns} \cdot NS_i(\psi) \right) \qquad \text{Equation 6.8}$$

Where:

PI :	Performance Index
$D_i(\psi)$:	Average delay per vehicle for phase group *i*, in s/veh
$NS_i(\psi)$:	Average number of stops per vehicle for phase group *i*, in stops/veh
W_d :	Weighted factor for average delay
W_{ns}:	Weighted factor for average number of stops
$INTX$:	Total number of intersections in the traffic network

For this research, W_d and W_{ns} are selected to be 1.0 and 10.0 respectively. These values were adopted from signal optimizations carried out using the SYNCHRO 6 traffic optimization software [52]. These values were selected so that the number of stops and average control delay are within the same scale, and to ensure that the GA will not be biased toward minimizing delay over number of stops.

The following subsections discussing in details the formulations used for estimating delays and the number of stops incurred by vehicles.

Average Control Delay Estimate

The HCM 2000 [55, 56] approach is used to calculate average control delay per phase group per intersection. For a given intersection, the control delay is calculated using the following formula:

$$D_n = d_{uni} \times PF + d_{inc} + d_{res} \quad \text{.......... Equation 6.9}$$

Where:

D_n: The average control delay per vehicle for phase group *n*, in seconds

PF: Uniform delay adjustment factor for quality of progression. According to the HCM 2000 [55] guidelines, and for the purpose of this research, a value of 1.0 is chosen

d_{uni}: Uniform delay assuming uniform arrival, in seconds

d_{inc}: Incremental delay due to traffic stochastic effects, assuming no residual demand at the beginning of the start analysis period, in seconds

d_{res}: Residual demand delay caused by the presence of queued vehicles at the beginning of the analysis period, in seconds

Uniform Delay

The uniform delay is defined as the delay occurred in a lane group if vehicles arrive with a uniform distribution and saturation does not occur in any cycle [56]. For each intersection, the uniform delay for a specific lane group is defined as:

$$d_{uni} = \sum_{i=1}^{\Phi} 0.50\,CL \frac{(1 - g_i / CL)^2}{1 - (g_i / CL)[\min(v / c_i, 1)]} \quad \text{.................................} \quad \text{Equation 6.10}$$

Where:

v / c_i :	Volume to capacity ratio for phase i
g_i :	Effective green duration for phase i, in seconds
Φ :	Total number of phases for each cycle

Incremental Delay

The incremental delay takes into consideration that the arrival is random, not uniform, and some cycles might overflow [56]. Mathematically, the incremental delay is expressed as follows in the HCM 2000:

$$d_{inc} = \sum_{i=1}^{\Phi} 900T \left[(v / c_i - 1) + \sqrt{(X_i - 1)^2 + \frac{8 k_i I_i v / c_i}{c_i T}} \right] \quad \text{..................} \quad \text{Equation 6.11}$$

Where:

T :	Analysis period duration, in hours
k_i :	Incremental delay factor based on controller settings. As per HCM 2000 guidelines [55], a value of 0.50 is assigned to this factor

I_i: Upstream filtering metering adjustment factor accounting for the effect of filtered arrival from the upstream. As per HCM 2000 recommendations [55], a value of 1.00 is assigned to this parameter in the current research

Residual Delay

The residual delay occurs as a result of an initial unmet demand at the beginning of the analysis period. For a given intersection, the calculation of residual delay done as per Equation 6.12. [56].

$$d_{res} = \sum_{i=1}^{\Phi} \frac{1800 Q_{bi}(1-u_i)t_i}{c_i T} \quad \text{Equation 6.12}$$

Where:

Q_{bi}: Initial unmet demand at the start of period *T* for phase *i*

t_i: Duration of unmet demand in *T* for phase *i*, (see Equation 6.13)

u_i: Delay parameter for lane group *i*, (Equation 6.14)

$$t_i = \begin{cases} 0.0 & IF \quad Q_{bi} = 0.0 \\ \min\left[T_i, \dfrac{Q_{bi}}{c_i(1-\min(1-X_i))}\right] & IF \quad Q_{bi} \neq 0.0 \end{cases} \quad \text{Equation 6.13}$$

$$u_i = \begin{cases} 0.0 & IF \quad t_i < T \\ 1 - \dfrac{Q_{bi}}{c_i(1-\min(1-X_i))} & IF \quad t_i \geq T \end{cases} \quad \text{Equation 6.14}$$

Number of Stops Estimation

In this research, the number of stops is defined as the number of vehicles that arrive at a downstream intersection during either the red interval or when queue exists [45]. For simplicity, some assumptions are made in the execution of the calculations:

- Only unsaturated traffic conditions are considered, (i.e., only situations with traffic demand less that the capacity for each cycle).
- All traffic demand is met during the cycle, with no overflow from the previous cycle to the next.

Although the above assumptions seem to simplify the estimation of the number of stops, these assumptions have been shown to fit well with simulation experiments [57].

For a given phase, the time required to dissipate a queue is a function of the red interval duration, arrival rate, and saturation rate. The queue dissipation time can be expressed as follows when considering average arrival, departure and saturation flow rates [45]:

$$TQ_i = \frac{S_i r_i}{S_i - \lambda_i} \qquad \text{Equation 6.15}$$

Where:

TQ_i: Queue duration for phase *i*, in seconds

r_i: Red interval for phase *i*, in seconds

S_i: Saturation flow for phase *i*, in veh/s

λ_i: Vehicle arrival flow for phase *i*, in veh/s

The queue dissipation time represents the portion of the green interval that effectively acts as a red signal for approaching traffic. The number of stops attributed to arrivals while a queue is present can then be calculated using Equation 6.16. [45].

$$NS_n = \sum_{i=1}^{\Phi} \lambda_i \times TQ_i \quad \text{......... Equation 6.16}$$

Where:

NS_n : Number of stops during phase n, in number of stops/veh

6.3.2. Transit Travel Time (TTT)

Figure 6.2 illustrates the trajectory of a transit vehicle traveling through its route in the network. Once a transit vehicle is detected at an intersection, it leaves the intersection and starts traveling on the segment linking its current intersection to the next one. If a bus stop is present between the two intersections, as in Figure 6.2, the vehicle may then temporarily stop for boarding and alighting passengers, if needed. Once the transit vehicle is done with dwelling, it resumes travel toward the downstream intersection, where it might stop again for a while waiting for any existing queue to clear, or for green time to start. Based on the above elements, Transit travel time, (TTT) can be expressed using Equation 6.17:

$$TTT = \frac{\overline{L}}{V_b} + \sum_{stp=b}^{PT} dwl_{stp} + \sum_{i=j}^{J} TD_i \quad \text{......... Equation 6.17}$$

Where:

TTT : Estimated transit travel time from a transit current location to the end of the network (i.e., the target destination), in seconds

$\overline{L}$: Distance between the transit current location and the end of the network, in ft

V_b : Transit desired speed, in ft/s

dwl_{stp} : The predicted time elapsed between the moment a transit vehicle enters the bus stop and the moment the bus will be ready to leave the bus stop *stp* and merge with traffic, which includes dwell time (will be discussed in detail in Chapter 7), in seconds

stp : Bus stop ID

b : The first coming bus stop meets the transit vehicle

PT : All the bus stops in front of a transit vehicle along its route

TD_i : Expected delay for a transit traveling through intersection i

j : The first downstream intersection will meet a transit vehicle

J : The remaining intersections that a transit vehicle needs to travel through

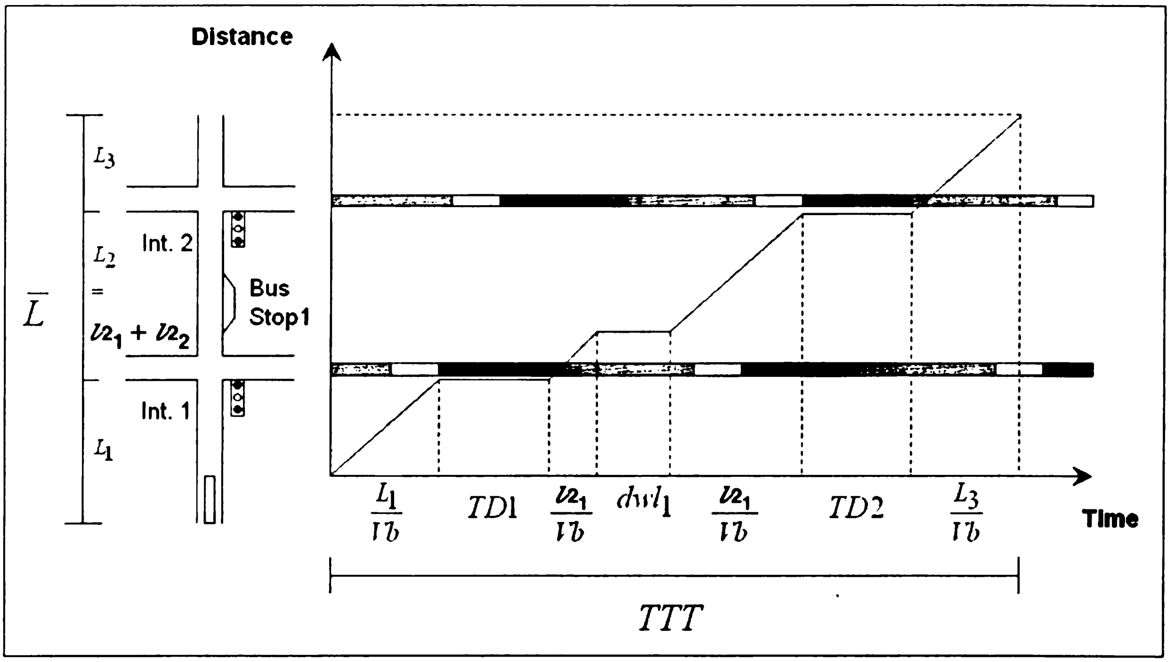


Figure 6.2 Time space diagram for transit vehicle traveling throughout the system

The simulation environment has the ability to provide the traffic controller with estimates for the initial queue at each intersection. In reality, information about the number of queued vehicles at any given time can be derived by individual signal controllers by compiling vehicle detections on intersection approaches. This information could then be communicated to adjacent signal controller if the controllers have the capability to communicate with each other, either through a local direct communication link or communication with a central traffic control center.

The first term in the equation predicting transit travel time is computed by simply dividing the distance remaining for the transit vehicles to leave the network from its current location by the vehicle's desired speed. If the transit vehicle experiences no delay, and does not dwell at any bus stop, the transit travel time would then equals the distance to travel divided by a transit vehicle desired speed.

An ANN model is used to predict the transit dwell time parameter, dwl_b, based on traffic, signal, and transit information. The prediction of this parameter will be discussed in Chapter 7. Note that this term represents the predicted time that a bus might take from the moment it enters a bus stop to the moment the bus is ready to leave. If the bus stop is a near side bus stop, the time the bus is ready to leave would then represent the stop line projected arrival.

The last term required to estimate transit travel time is TD_i. For a given analysis period, the average delay is calculated as the summation of the total delay divided by number of vehicles that pass within this period, which means that not all the vehicles are subjected to delay. Therefore, it is important to separate the transit vehicle from the traffic stream, and look at its expected delay individually.

If a vehicle arrives at the beginning of the red interval, it experiences the maximum delay, since it remains still during the entire red interval. Vehicles arriving later during the red are also forced to stop; however, their delay would not be as high as vehicles arriving at the beginning of the red. When the signal turns green, the first vehicle starts moving, followed by the second vehicle, until the queue is dissipated. Once the queue is dissipated, any arrival vehicle does not have to stop, and therefore, does not face any delay leaving the intersection.

The delay behavior in this case is linear. The maximum delay is when a vehicle arrives at the beginning of the red interval, and the zero delay is when the vehicle arrives after the downstream queue is dissipated. Therefore the expected delay can be calculated using Equation 6.18:

$$TDi = \begin{cases} 0.0 & IF \quad TQ_i < Arrival < CL \\ \dfrac{2Dn \times CL(1 - Arr_i)}{TQ_i^2} & IF \quad 0 \le Arrival \le TQ_i \end{cases} \quad \text{......... Equation 6.18}$$

Where:

Arr_i: The stop-line projected arrival time at intersection *i* relative to the cycle local timing given that no is queue presents, in seconds

TD_i : Expected transit delay at intersection *I*, in seconds

6.3.3. Schedule Adherence Penalty (SCH)

Buses are expected to arrive at each bus stop following a given schedule. However, due to random traffic behavior, it is possible that individual buses will arrive earlier or later than expected. In order to accommodate early or late arrivals, a penalty is introduced as a function of the transit schedule and the anticipated bus stop arrival. The schedule adherence penalty is added to the cost function for any solution in the search space for which the absolute difference between the scheduled and projected arrival time of a vehicle at a specific bus stop exceeds zero. For a specific bus stop, this penalty can be expressed by Equation 6.19:

$$SCH_{stp,b} = \left| SchArr_{stp,b} - ExpArr_{stp,b} \right| \quad \text{................................ Equation 6.19}$$

Where:

$SCH_{stp,b}$: Bus *b* penalty due to early or late arrival at bus stop *stp* arrival according to the schedule

$SchArr_{stp,b}$: Bus *b* scheduled arrival at bus stop *stp*, in seconds

$ExpArr_{stp,b}$: Bus b expected arrival at bus stop stp, in seconds

The transit vehicle expected arrival is computed by utilizing Equation 6.17, as presented in the following equation:

$$ExpArr = \frac{\overline{L}}{V_b} + \sum_{stp=b}^{PT} dwl_{stp} + \sum_{i=j}^{J} TD_i \qquad \text{Equation 6.20}$$

Where:

$\overline{L}$: Distance between the transit location and the target bus stop, *stp* when the optimization process is called, in ft

j : The first intersection to meet the transit vehicle before reaching the target bus stop, if any

J : The last intersection to meet the transit vehicle before reaching the target bus stop, if any

b : The first bus stop to meet the transit vehicle while approaching the target bus stop, if any

PT : The last bus stop to meet the transit vehicle while approaching the target bus stop, if any

stp : Bus stop ID

For optimized signal timings in a network consisting of multiple links with multiple bus stops, the cost function will be penalized by the sum of all the penalties applied at all bus stops in the network, for all the buses currently in the network. A mathematical expression for this calculation can be written as:

$$SCH = \sum_{\forall stp \in S}^{S} \sum_{\forall b \in B}^{B} SCH_{stp,b} \qquad \text{Equation 6.21}$$

Where:

SCH : Schedule adherence penalty, in seconds

B : Number of transit vehicles that are currently in the network

S : Number of bus stops a transit vehicle did not pass yet along its route.

6.4. Optimization and Simulation Integration

Previous sections of this chapter represented the cost function that is to be fed to the controller. As discussed in Chapter 3, the proposed algorithm is set to accommodate problems associated with dwell time variability, implementing TSP within a coordinated signalized network, and the need to handle simultaneous priority requests. Furthermore, the D-SPORT controller needs to be able to optimize general traffic movements, in cases in which transit vehicle may or may not be present in the network.

The flowchart shown in Figure 6.3, describes the integration of the optimizer within the simulator by using the VISSIM-COM [58] interface. This interface allows communicating to the simulation environment information to and from an external controller. This integration process will be discussed in details in Chapter 8.

In the illustrated flowchart, the VISSIM-COM interface triggers the start of the simulation process. The simulation moves forward until the time for calling the optimizer has come, which is a function of signal controller's cycle lengths and offsets. Once the controller decides to call the optimizer, information such as traffic data, transit data, and controller status is passed to the optimizer. The GA optimization platform, which is the tool that is used to perform the optimization, then finds a near-optimal solution by minimizing the D-SPORT cost function, and passes the signal timings back to the controllers through the VISSIM-COM interface. This process is iterated until the 2-hr simulation duration has elapsed.

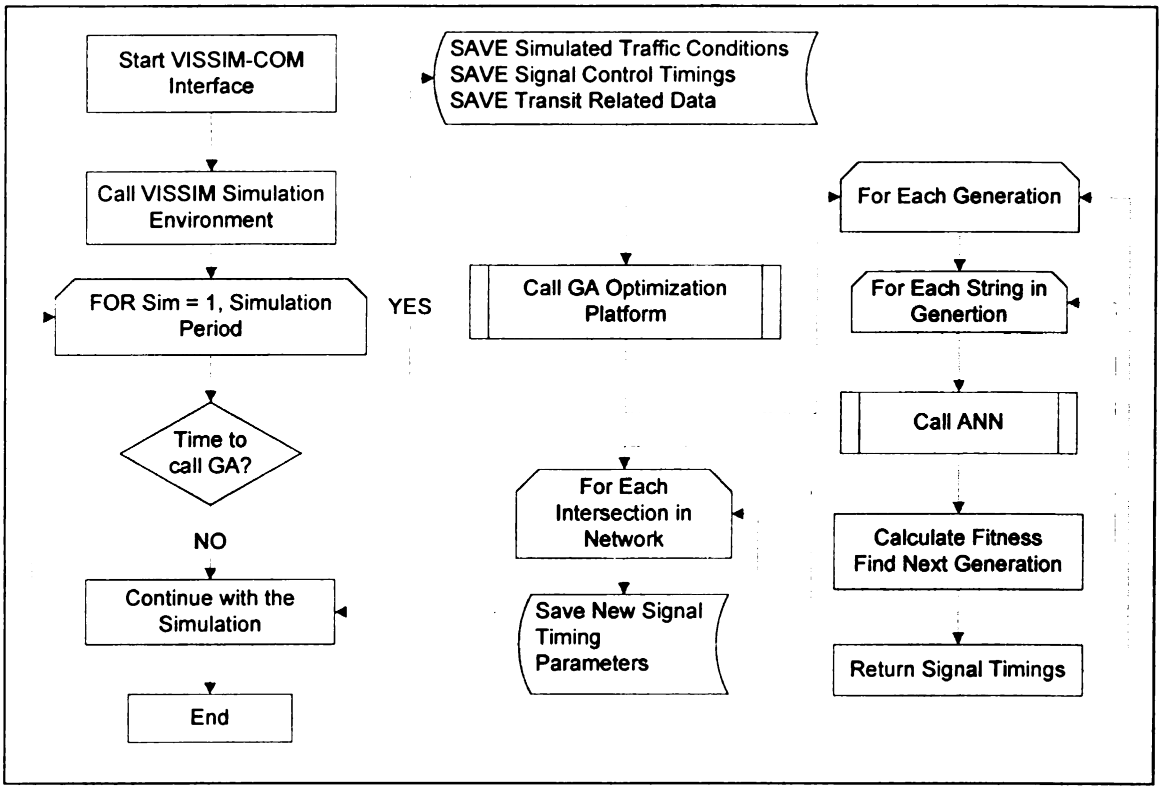


Figure 6.3 Flowchart for optimization process and simulation environment integration

CHAPTER VII.

TRANSIT TRAVEL TIME PREDICTION MODEL

This Chapter presents the development of an ANN-based prediction tool for providing information about the expected arrival time of buses at intersections where dwell activities of a near-side bus stop may create uncertainties in the predictions. As detailed in Chapter 6, an ANN prediction model is required for determining the transit travel time term in the D-SPORT cost function. The ANN prediction model presented in this chapter takes advantage of the properties of artificial neural networks. To ensure that the model could be applied in real-world settings, its development is also based on constraints stipulating that the travel time predictions are to be based on information that is available to traffic signal controllers, traffic information systems, or that can be obtained from current sensing technologies. The following sections successively describe the experimental setup, the initial list of model parameters that were identified for possible inclusion in the prediction tool, the process by which data simulating the bus arrival processes were generated for the development of the prediction model, the process by which a final selection of model parameters was conducted, the training of the proposed neural network model, and the evaluation of the ANN model performance.

7.1. Experimental Setup

The test network used for the model development is based on the network that was used in Chapter 5 for the evaluation of dwell time variability impacts on TSP operations. This network is illustrated in Figure 5.1. It is used in this part of the research to generate the datasets for building and training the proposed ANN model. The only changes that were made to the simulation model described in Chapter 5 are related to the traffic signal control operations, traffic demand, and transit service demand. These input parameters were adjusted to allow the execution of simulations covering a wide range of possible combinations of traffic demand, transit demand and traffic signal control operations. Below is a brief summary of the key features of the simulation modeling used in this portion of the research.

Traffic Signal Control

Each intersection is controlled by a 90-s two-phase fixed-time signal operation providing 40-s of effective green time on each approach. A 20-s minimum green is further imposed on each phase to ensure that priority actions do not unduly restrict pedestrian crossing times. Based on the traffic demand assigned to each experiment, the signal coordination between the two intersections is optimized using SYNCHRO 6 traffic signal timings optimization software.

Traffic Demand

To remove truck effects, only passenger cars are simulated. Vehicles are generated at the upstream end of each entry link using a Poisson distribution and VISSIM's default vehicle characteristics. To allow testing in a range of traffic conditions, scenarios with eastbound (EB) and westbound (WB) volumes of 600, 800, 1000, 1200, 1400, and 1600 vph, and northbound/southbound volumes corresponding to 80% of the eastbound/westbound volumes, were considered. For each scenario, additional sub-scenarios were created by varying the proportions of right-turning and left-turning vehicles on the eastbound approach to Intersection 1 between 5% and 15%. Turning percentages on all other approaches remained fixed at 15% right turns and 5% left turns.

Transit Service

Buses are assumed to travel in the eastbound (EB) direction only and to service a single stop located in the right-turn bay of the eastbound approach to Intersection 1 with a frequency of one bus every 5-min, with a standard deviation of 20-s. While buses are generated to enter the network with a 20-s standard deviation and all share a desired 65 km/h (40 mph) travel speed, interactions with other vehicles and traffic signals create more fluctuations in the time each vehicle takes to reach the bus stop. To minimize the impacts of gap-seeking delays, all buses are modeled to have priority over the adjacent traffic when leaving the bus stop, which can be regulated in the field. While this facilitates their departure, it does not completely eliminate the need to seek gaps in the adjacent traffic, particularly when a queue exists in the adjacent lane.

Passenger demand is modeled by simulating the process of passenger arrivals at the bus stop. Passengers are modeled to arrive at an average rate varying between 25 and 200 passenger/h, with an increment of 25 passenger/hr. The generated passengers are assumed to enter the bus stops by following the Poisson distribution, since the simulation environment does not support other distributions for transit demand. For each bus, dwell time is then determined by assuming that 3-s are required to allow each passenger to board, with another 3-s for door opening and closing. Dwell time variations thus result from randomness in passenger arrivals at the stop and the effects of early or late transit arrivals.

Figure 7.1 summarizes the simulation experiments setup that are used to generate the data required to train the ANN model. As can be seen in the sketch, different traffic volumes are combined with different left and right turning volumes and different transit demand at the bus stop. Based on the various combinations considered, a total of 360 scenarios were built and each simulated for a 2-hr period.

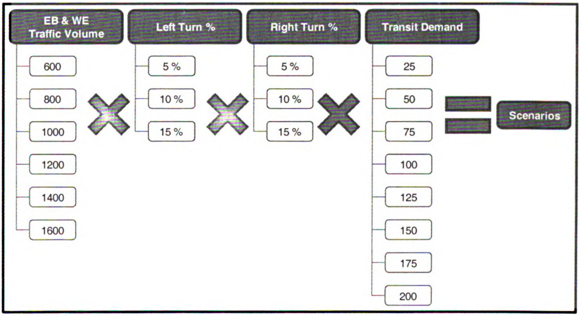


Figure 7.1 Experiments used to generate the datasets required for the ANN model development

7.2. Initial Candidate Model Parameters

Variables shown in Table 7.1 were initially selected as tentative input variables to the ANN model. The time elapsed between detection and the moment a bus is ready to leave the bus stop is the value to be predicted. All the selected variables quantify traffic conditions along the approach on which the bus is traveling. Most of the selected variables further represent average traffic conditions in 5-min sampling intervals. The information represented by these variables will not necessarily represent the average traffic conditions in the 5-min that preceded a detection, but rather the conditions in a 5-min window that may have ended anywhere between 0 and 5 min prior to the detection. Observations over a 5-min interval are used to reduce the sensitivity of the prediction model to short-term traffic variability and allow the model to focus on short-term trends rather than responding to short-lived random fluctuations, and to make the model more robust to handle various conditions. Also, a wider time interval is not used since collecting data over a wider period of time, says 15-min, would cause a loss of real-time data that are used to feed the dwelling activity prediction model.

Table 7.1 Variables considered for the ANN prediction model

Variable	Definition
	Input Variables
Pass ***(Persons/bus)***	Number of passenger in the bus
Load ***(Passenger/bus stop)***	Current number of passenger at the bus stop
THV ***(#veh)***	Number of through vehicles in previous 5-min interval
RTV ***(#veh)***	Number of right turn vehicles in previous 5-min interval
LTV ***(#veh)***	Number of left turn vehicles in previous 5-min interval
THTT ***(sec/veh)***	Travel time of through vehicles in previous 5-min interval
RTTT ***(sec/veh)***	Travel time of right turn vehicles in previous 5-min interval
LTTT ***(sec/veh)***	Travel time of left turn vehicles in previous 5-min interval
THD ***(sec/veh)***	Average delay of through vehicles in previous 5-min interval
RTD ***(sec/veh)***	Average delay of right turn vehicles in previous 5-min interval
LTD ***(sec/veh)***	Average delay of left turn vehicles in the previous 5-min interval
THQ ***(ft)***	Average queue of through vehicles in previous 5-min interval
RTQ ***(ft)***	Average queue of right turn vehicles in the previous 5-min interval
LTQ ***(ft)***	Average queue of left turn vehicles in previous 5-min interval
THQM ***(ft)***	Maximum queue of through vehicles in previous 5-min interval
RTQM ***(ft)***	Maximum queue of right turn vehicles in previous 5-min interval
LTQM ***(ft)***	Maximum queue of left turn vehicles in previous 5-min interval
LCyc ***(s)***	The current local cycle time when bus is detected
Late	Transit Lateness relative to the previous transit relative excluding average transit headway
Offset ***(s)***	The offset between the upstream intersection relative to the downstream intersection
	Predicted Variable
Arr ***(s)***	Time elapsed between bus detection and the moment the vehicle is ready to leave the bus stop

7.3. Data Generation

Similar to the simulations of Chapter 6, VISSIM is used to generate data for the development of the ANN model. Simulated data were used due to difficulties in collecting all the required data from real-world intersections, where typical existing detection technologies may not allow easy collection of many of the parameters being considered. This includes offsets, transit occupancies, and current local cycle. It is furthermore difficult to control experiments in real field settings. Contrary to real field experiments, simulation allows more flexibility in controlling the experimental setup and performing sensitivity analysis.

For each simulated bus, input variables along with the dependent variable were stored. These stored variables were used to generate the datasets for developing the ANN model. Once the datasets were generated, input and output variables were normalized to have their values only ranging between 0 and 1, as shown in Equation 7.1.

$$\hat{X}_i = \frac{X_i - X_{\min}}{MAX(X_i) - MIN(X_i)} \qquad \text{Equation 7.1}$$

Where:

$\hat{X}_i$: The normalized variable X_i

X_i: The variable i value

$MIN(X_i)$: The minimum value for input X_i in the collected data

$MAX(X_i)$: The maximum value for input X_i in the collected data

The transit time consumed at the bus stop (the dependent variable) was taken as the time elapsed between the moment a transit enters a bus stop and the moment the transit has closed its door and is ready to move. This includes the deceleration until complete stops, the dwell time, a fixed 3-s for door opening, and another 3-s for door closing. Note that the delay resulting from having the transit vehicles wait for a gap in the adjacent traffic to enter the intersection (merge can occur in the intersection) is not part of the predicted variable. This is because such delays are more closely associated with traffic interactions than dwelling activities. Furthermore, non-transit vehicles are modeled to yield for transit vehicles when they merge back in the traffic stream, which would reduces the significance of gap-seeking delay impact on traffic behavior is minimal. Nonetheless, this type of delay is accounted for in the transit travel time sub-cost function defined in Equation 6.17 through the TD_i term.

For non-dwelling buses, and for buses arriving during the normal green, the predicted dwell time is the time elapsed between activation of the check-in and check-out detectors. For buses arriving at the downstream intersection during the red the time, the predicted time is the time elapsed between activation of the check-in detection and the moment the bus is fully stopped.

The combination of 6 traffic volumes, 3 right turning percentage, 3 left turning percentages, and 8 boarding rates resulted in 432 scenarios. For each scenario, the consideration of a two-hour simulation period resulted in the generation of 23 buses. Out of the 9936 potential datasets that could be produced, only 8871 were generated, as traffic interactions prevented some buses from exiting the network before the end of the simulation.

To ensure the validity of the ANN training process, three independent groups of datasets were created by randomly assigning each of the 8871 datasets produced to a specific group:

- Training datasets: 5000 datasets used to train the ANN in predicting the dependent variable.

- Validation datasets: 2000 datasets used to monitor the performance of the ANN during the training process and to keep the network from over-fitting, which occurs when the ANN model may fit the noisy data, and not only the signal, which reduces the ANN model ability to be generalized beyond the training datasets. The use of an independent group of datasets for validation supervises the training process and increases the prediction ability of an ANN for cases not seen when submitting training datasets.
- Testing datasets: 1871 datasets used to examine the ability of ANN to predict the dependent variable in cases independent from the training and validation datasets.

7.4. Selection of ANN Model Parameters

To help with the final selection of model variables, the training datasets were used to generate the correlation matrix of Table 7.2. This matrix shows the result of correlation analyses conducted between each pair of parameters. These analyses led to the rejection of the variables shown in the shaded columns due to their high colinearity with other variables. The rejected variables are travel time for through, right, and left movements, (THTT, RTTT, LTTT), and the maximum queue for left, through, and right movements, (LTQM, THQM, RTQM). Although some of the retained variables also appear correlated to others, these were not excluded from the training process since their inclusion in the ANN model was found to improve its prediction performance, mainly by reducing the root mean square error (RMSE) and improving the coefficient of determination (R^2). A similar argument was also invoked to retain some variables with low correlation with the travel time variable to predict.

Table 7.2 Dependent and independent variables correlation matrix

Arr	Pass	Load	THV	RTV	LTV	THTT	RTTT	LTTT	THD	RTD	LTD	LTQ	TRQ	RTQ	LTQM	THQM	RTQM	LCyc	Late	Offset
1.000	0.518	0.96	-0.014	-0.008	-0.036	-0.024	0.153	-0.114	-0.024	0.156	-0.114	-0.133	-0.133	-0.124	-0.118	-0.117	-0.1.00	0.013	-0.019	0.002
	1.000	0.383	0.034	-0.012	-0.017	-0.014	0.104	-0.133	-0.013	0.107	-0.133	-0.138	-0.138	-0.13	-0.124	-0.125	-0.095	-0.023	0.105	0.014
		1.000	-0.019	-0.009	-0.036	-0.013	0.153	-0.085	-0.013	0.156	-0.086	-0.097	-0.097	-0.088	-0.086	-0.084	-0.071	0.026	-0.044	0.001
			1.000	-0.122	0.173	0.257	0.194	0.267	0.249	0.191	0.266	0.243	0.262	0.222	0.305	0.337	0.296	0.094	0.098	0.003
				1.000	0.267	0.281	0.146	0.22	0.279	0.13	0.219	0.17	0.163	0.17	0.216	0.207	0.229	-0.003	0.008	0.004
					1.000	0.343	0.129	0.422	0.338	0.123	0.419	0.331	0.307	0.291	0.409	0.354	0.343	0.042	-0.01	0.027
						1.000	0.176	0.443	1.00	0.171	0.443	0.353	0.354	0.33	0.408	0.408	0.399	-0.21	0.026	0.014
							1.000	0.167	0.173	0.993	0.167	0.152	0.156	0.155	0.19	0.202	0.223	0.056	0.092	0.002
								1.000	0.435	0.164	1.00	0.728	0.712	0.691	0.778	0.754	0.742	0.035	-0.207	0.005
									1.000	0.168	0.435	0.345	0.346	0.323	0.4	0.399	0.391	-0.214	0.028	0.014
										1.000	0.163	0.151	0.154	0.154	0.187	0.199	0.221	0.057	0.093	0.002
											1.000	0.729	0.713	0.692	0.778	0.755	0.743	0.035	-0.207	0.005
												1.000	0.998	0.991	0.976	0.979	0.971	0.059	-0.14	-0.002
													1.000	0.992	0.972	0.98	0.972	0.06	-0.139	-0.003
														1.000	0.959	0.966	0.972	0.058	-0.138	-0.003
															1.000	0.989	0.978	0.065	-0.134	-0.001
																1.000	0.986	0.069	-0.131	-0.003
																	1.000	0.065	-0.124	-0.002
																		1.000	-0.291	0.005
																			1.000	-0.012
																				1.000

Highlighted columns were neglected when the final ANN model is developed

7.5. ANN Prediction Model Training

A feedforward backpropagation learning algorithm (BP) was used to train different ANN models. This is because BP learning algorithms are currently the most commonly used learning tools for multi-layer ANNs [59]. The underlying theory behind BP algorithms is the utilization of non-linear perceptron-like systems. Training is achieved by minimizing the mean-squared-error (MSE) between the actual and desired outputs of the multi-layer feedforward perceptron.

Equation 7.2 shows the objective function that was used for ANN training. As explained above, the executed training sought to minimize the MSE between the predicted and observed dwell time values.

$$\varepsilon = \frac{1}{n_{tr}} \sum_{r=1}^{n} (a_r - p_r)^2 \qquad \text{................................ Equation 7.2}$$

Where:

ε:	MSE over entire sets of neurons
a_r:	Actual value of the dependent variable for dataset *r*
p_r:	Predicted value of the dependent variable for dataset *r*
n_{tr}:	Number of training datasets

Learning takes place through the repeated modification of the network weights based on their relative contribution to the error term. Specifically, the derivative of the error function with respect to any weight in the network is computed and used to adapt that particular weight to meet the objective of the function to minimize the error value, as shown in the following equation:

$$\Delta W_{ij} = \eta \frac{\partial \varepsilon}{\partial W_{ij}} \quad \text{...} \quad \text{Equation 7.3}$$

Where:

ΔW_{ij} : Change of weight between nodes i and j

η : Learning rate, which is the rate at which the weights between neurons are adjusted

There are two possible learning methods to train an ANN, the batch training and the incremental learning. In the batch learning, the weight adjustments are accumulated over the whole training sets, and updated at the end of each iteration in the ANN learning process, which is also known as an epoch. In the incremental training, the weights are instead adjusted for every dataset (pattern) before the next pattern is considered. For this research, batch training is used since it could result in a high precision data mapping [60, 61].

Batch learning was used to train different ANN architectures for the development of an ANN model able to accurately predict the desired output variable. The tested ANNs differ by the number of hidden layers and neurons, the transfer functions, and the learning rate used. Based on results of the learning process, it was determined that 500 epochs, which is the number of cycles where the training process loops over the training dataset, was sufficient to build an ANN model, as all the examined architectures were trained before reaching the 500 epochs.

The different ANN architectures to test were developed using MATLAB 6, which is a numerical computing environment and programming language [62]. MATLAB 6 contains a neural network toolbox that allows users to build, train, and test customized ANN architectures. The MATLAB code that is used to develop the ANN model can be found in Appendix A.

As shown in Table 7.3, the different architectures involved different numbers of processing layers (3 or 4), hidden layers (1 or 2) and hidden neurons (8 to 23). The learning rate was changed between 0.10 and 0.50. Three types of transfer functions were further used: *logsigmoidal, tansismoidal* or *purelinas*. These are described in Equation 7.4, Equation 7.5, and Equation 7.6 respectively.

$$logsig(X_i)=\frac{1}{1+e^{-X_i}} \qquad \text{Equation 7.4}$$

$$tansig(X_i)=\frac{2}{1+e^{-2X_i}}-1 \qquad \text{Equation 7.5}$$

$$purelin(X_i)=X_i \qquad \text{Equation 7.6}$$

Where X_i is the input variable.

For each ANN model, the coefficient of determination R^2 was obtained by comparing travel time predictions with the simulated values. This parameter assesses how much of the observed travel time variance is explained by the model considered.

The best architecture was found to be ANN Model 10. This model exhibits the highest R^2 across the training, validation and testing datasets. This model has one hidden layer with 8 neurons and a learning rate of 0.10. The *logsigmoidal* transfer function was used between the input and the hidden layer, while the *purelin* transfer function was applied between the hidden and the output layers. While other models showed equivalently high R^2 values, they had longer training times. The validity of the selected model is demonstrated by the plots of Figure 7.2, which compare predicted and simulated (observed) travel times for the various datasets considered. As shown, the comparisons produced clear linear relationships between the predicted and simulated data, in addition to a relatively low RMSE.

However, there are a few points that do not follow the expected linear behavior. This is because the model did not consider all the factors can be associated with the dependent variable prediction, such as driver's behavior, vehicle characteristics, and lane changing models. All these factors had some interaction effects in the simulations from which the data used to develop the ANN model were retrieved. While excluding the interaction between these factors and the predicted variable could potentially lower the prediction accuracy, these factors are difficult to measure in reality, and thus, difficult to incorporate into the prediction model.

Table 7.3 **Performance evaluation of various neural network configurations**

Model ID	# Layer *	Learning Rate	ANN Structure**	Transfer Function Type			R^2		
				Input ↓ Hidden	Hidden ↓ Hidden	Hidden ↓ Output	Train	Validate	Test
1	4	0.50	14-(15-8)-1	logsig	tansig	purelin	0.972	0.956	0.966
2	4	0.50	14-(15-5)-1	logsig	tansig	purelin	0.970	0.958	0.968
3	4	0.50	14-(10)-1	logsig	tansig	purelin	0.972	0.958	0.966
4	3	0.50	14-(8)-1	logsig	-------	purelin	0.970	0.953	0.960
5	3	0.50	14-(8)-1	tansig	-------	purelin	0.970	0.956	0.964
6	4	0.50	14-(8-2)-1	logsig	logsig	purelin	0.958	0.949	0.955
7	4	0.50	14-(8-2)-1	logsig	purelin	purelin	0.960	0.943	0.953
8	3	0.25	14-(8)-1	logsig	-------	purelin	0.972	0.956	0.964
9	3	0.15	14-(8)-1	logsig	-------	purelin	0.970	0.955	0.962
10	***3***	***0.10***	***14-(8)-1***	***logsig***	***------***	***purelin***	***0.994***	***0.994***	***0.998***

* Number of layers means the number of input, hidden, and output layers

** ANN 14 - (10-5) - 1 represents the structure of a neural network has 14 inputs, two hidden layers with 10 and 5 hidden nodes respectively, and 1 output

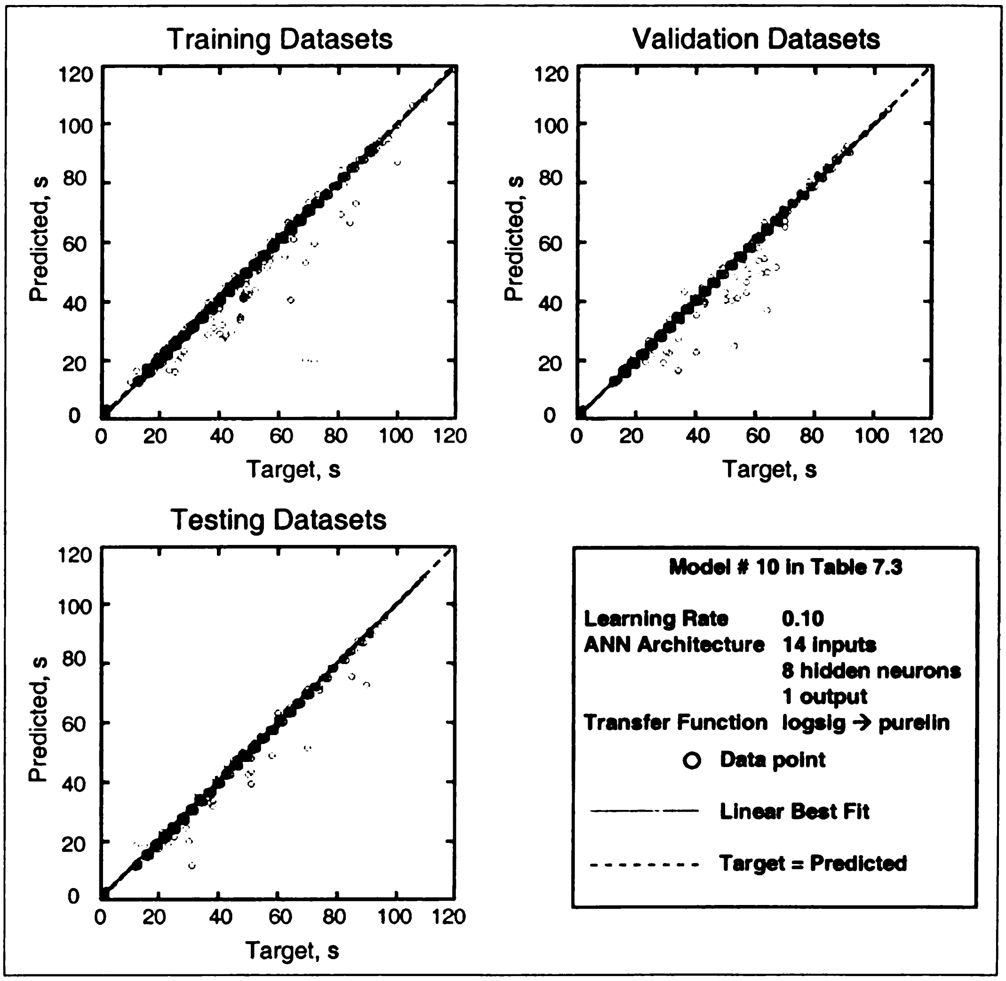


Figure 7.2 **Regression analysis of predicted against observed travel times**

7.6. Enhanced Active TSP Performance

Following the development of the ANN model, the research focused on integrating active TSP system and the ANN model within the control system, and examining the performance of the enhanced TSP system using the VISSIM simulation environment. The communication between the enhanced TSP signal control system and the simulation environment is done by using the VISSIM-COM interface according to the flowchart of Figure 6.3 presented earlier.

Within the integrated model, input variables corresponding to the retained variables listed in Table 7.2 (variables not highlighted) are sent to the ANN model each time the upstream check-in detector is activated by a bus. Upon reception of the information, the prediction model uses it to determine when the approaching bus is most likely to be ready to leave the bus stop and sends this prediction to the signal controller, which in turn decides if a priority action is needed, and what TSP action should be granted.

The appropriate priority action is determined at each decision point based on the following rules:

- No action is taken for buses projected to arrive during the green.
- For buses arriving soon after the end of green, a green extension is granted if the projected arrival is within the maximum allowed extension. Extensions are awarded in increments of 1 s, up to a maximum of 20 s.
- Green extensions are terminated when a bus crosses the check-out detector.

- Early recalls are implemented for buses projected to arrive during the red beyond the maximum green extension window but close enough to the end of the red to be within the limit of permitted alterations.
- Any phase cannot be shortened beyond its specified minimum green.
- Each bus is allowed only one priority request. A bus failing to enter the intersection in time to benefit from a green extension will thus not be considered for an early recall.
- Buses arriving during the normal green are not considered as they do not require priority.
- Buses requesting priority and entering the intersection during a provided green extension (GE) or early recall (RT) are counted as successes.
- A bus requesting a green extension and arriving too late to benefit from it is considered as a failure.
- A bus requesting a green extension but arriving earlier than expected and for which a priority request cancellation can be issued before the extension is implemented is considered as a success since the system was able to adjust to the early arrival.
- The awarding of early green recalls for buses that end up arriving during the normal green is considered as a success as reduced delays may still result from early queue dissipation.

To assess any incremental benefits that may be associated with the use of the ANN model, the performance of the enhanced active TSP that integrates the ANN model within the controller is first compared to the performance of a fixed-time traffic signal control system offering no preferential treatments. A second series of comparisons is then made with a controller making a standard TSP decisions by considering only an average dwell time and an average travel time between the check-in and check-out detectors.

Impacts of priority actions on transit operations are evaluated by comparing the resulting average bus delays and bus travel times. Traffic impacts are for their part assessed by comparing average traffic delay along the east/west and north/south bounds. For buses, the delays estimated by VISSIM includes bus stop clearance times but exclude all dwell times as dwelling activities are viewed as normal transit activities. Another performance measure examines system efficiency by calculating the ratio of buses having benefited from the system to the total number of buses having requested a priority action. Successes in responding to priority requests are assessed according to the rules explained earlier in this Section.

Four evaluation scenarios are finally considered by assuming volume-to-capacity ratios of either 50% or 85% and average boarding rates of either 30 or 180 passenger/hr. Similar to previous simulations, each scenario was run for two hours to allow observation of system performance over a number of buses. For each scenario, 10 replications with different random seeds were also executed to reduce the impacts of simulation stochastic effects on the evaluation results.

Figure 7.3 compares the predicted times required by buses to travel from the check-in detector to the intersection stop line to the actual travel times that were observed in the simulations. As can be observed, a reasonable agreement exists between the observed and predicted travel times, which is an indication of the general validity of the developed

ANN mode. Moreover, An R^2 of 0.8832 is achieved, which means that the ANN model is explaining over 80% of the variance of the observed results. However, there are some points that appear to be out of order. This observation can be explained by the fact that there are many other factors that can affect the accuracy of travel time prediction that need further investigation, such as lane changing models and other simulation modeling effects.

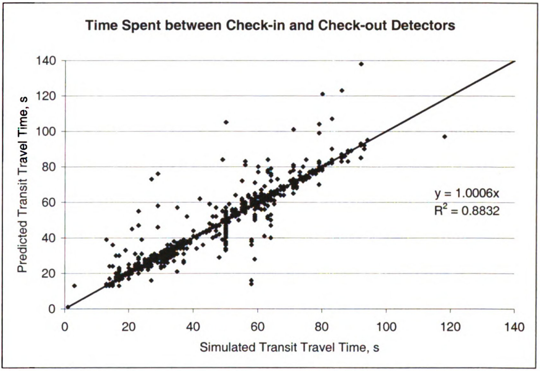


Figure 7.3 Comparison of simulated and predicted travel times between check-in detector and stop line for the evaluation scenarios

Table 7.4 further summarizes the measures of effectiveness (MOE's) obtained for the four scenarios considered. As expected, transit vehicles benefit from the provision of priority, with 95% statistically significant delay reductions of 7.0 to 13.0 seconds, or 35% to 50 % compared to the other two examined signal control systems. The simulation results further indicate that use of the ANN prediction model generally leads to an additional 1.0 to 3.0 seconds reduction in average bus delay and 4.0 to 10.0 seconds in average travel time when compared to a priority system considering only average dwell and travel times. These results are attributed to the better implementation of priority actions to the needs of approaching vehicles that result from the ANN model's enhanced travel time prediction capabilities.

Table 7.4 Measures of effectiveness for tested controllers

Scenario	Measure of Effectiveness	(1) No TSP	(2) Standard TSP	Change (1)→ (2) (%)	(3) Enhanced TSP	Change (1)→ (3) (%)	Change (2)→ (3) (%)
v/c: 0.50 30 passenger/hr	Bus Delay (s)	31.26	18.16	*-41.9 **	15.88	*-49.2 **	*-12.6 **
	Bus Travel Time (s)	50.29	36.60	*-27.2 **	31.36	*-41.6 **	*-19.8 **
	TSP Efficiency (%)	-----	97.00	-----	99.00	-----	-----
	Traffic Delay (s)	13.60	11.77	*-13.5 **	11.42	*-16.1 **	*-3.0*
	Crossing Delay (s)	14.35	17.07	*19 **	16.12	*12.3 **	*-5.6 **
v/c: 0.50 180 passenger /hr	Bus Delay (s)	19.02	14.26	*-25.0 **	12.44	*-34.6 **	*-12.8 **
	Bus Travel Time (s)	77.74	72.77	*-6.4 **	63.42	*-23 **	*-17.7 **
	TSP Efficiency (%)	-----	49.00	-----	88.00	-----	-----
	Traffic Delay (s)	13.66	12.94	*-5.3 **	12.67	*-7.2 **	*-2.1*
	Crossing Delay (s)	14.35	15.25	*6.3 **	14.74	*2.7 ^*	*-3.3 **
v/c: 0.85 30 passenger /hr	Bus Delay (s)	30.75	18.25	*-40.7 **	15.85	*-48.5 **	*-13.2 **
	Bus Travel Time (s)	49.81	36.72	*-26.3 **	30.44	*-43.5 **	*-23.3 **
	TSP Efficiency (%)	-----	98.00	-----	100.00	-----	-----
	Traffic Delay (s)	15.86	14.18	*-10.6 **	14.16	*-10.7 **	*-0.1*
	Crossing Delay (s)	12.58	14.53	*15.5 **	14.20	*12.9 **	*-2.3 ^*
v/c: 0.85 180 passenger /hr	Bus Delay (s)	19.86	13.45	*-32.3 **	12.38	*-37.7 **	*-8.0 **
	Bus Travel Time (s)	78.83	72.11	*-8.5 **	66.72	*-15.4 **	*-7.5 **
	TSP Efficiency (%)	-----	64.00	-----	82.00	-----	-----
	Traffic Delay (s)	15.83	14.90	*-5.9 **	15.23	*-3.8 **	*2.2*
	Crossing Delay (s)	12.65	13.69	*8.2 **	13.35	*5.5 **	*-2.5*

* Statistically significant at 99% confidence level compared to Case 1, No TSP

^ Statistically significant at 95% confidence level compared to Case 1, No TSP

Table 7.4 also shows that buses tend to benefit more from the provision of modified TSP controller when the transit demand is low, which results in a lower dwell time. This is due to a reduction in the accuracy of predictions with higher boarding rates. Under low boarding rates, the standard deviation in the distribution of passenger arrivals could be of a few persons only. This translates into small dwell time variations and a small window of potential arrival times at the downstream intersection. If the boarding rate is high, a similar standard deviation results in a higher variation in the number of persons to serve and, eventually, wider impacts on dwell time duration and the potential arrival window.

For a certain boarding rate, it is further observed that traffic congestion level only has a slight impact on transit benefits. This observation is expected, as buses are assumed to have priority to merge in the traffic stream when leaving the bus stop. While this priority helped the buses reentering traffic in the simulations, it did not completely eliminate associated delays, as buses still had to wait on average between 1.0 and 16.0 seconds to reenter traffic.

Results further show that system efficiency is the highest when the enhanced active TSP is implemented. This is due to the higher probability of successfully implementing priority actions from which the buses can fully benefit as a result of a higher accuracy in predicting bus arrival times.

Signal priority is finally observed to reduce arterial traffic delay between 4% and 16% and increase cross-street delay between 3% and 13% when ignoring the type of priority algorithm used. However, it can be observed that use of the enhanced TSP tends to produce slightly higher arterial delays and lower cross-street delays when compared to scenarios with no TSP or standard TSP treatments. In this case, a controller relying only on average values implements a certain number of signal timing alterations that may end up not used by approaching buses. Use of the ANN model enhances prediction capabilities then provides more opportunities to ignore priority requests that may end up not benefiting approaching buses. This translates into fewer green extensions and early recalls. Therefore, less effective green time is lost, less delays for the traffic traveling along the bus route, and less disruptions for the cross-street traffic is experienced in the traffic network.

7.7. Summary

An ANN prediction model is developed to predict the stop line arrival time of buses traveling on intersection approaches with near-side bus stops. Results of integrating the ANN model with active early green and green extension TSP algorithm shows an ability to improve the benefits that transit signal priority systems can provide to buses by enhancing the ability to correctly predict bus arrival times at intersections and reduce unnecessary signal timing adjustment. In the test scenarios, this enhanced prediction capability resulted in greater delay reductions for buses and lower traffic impacts than a priority system simply basing its decision on average dwell and travel times.

CHAPTER VIII.

D-SPORT FRAMEWORK INTEGRATION

This chapter describes the proposed optimization model that has been developed within the VISSIM simulation software for implementing TSP within coordinated networks of intersections. This development is based on the use of the VISSIM-COM interface to allow communications between the GA optimization process, the ANN prediction model, and the simulation environment.

8.1. Genetic Algorithm Platform

The first step in the development of the optimization process was to develop a GA-based optimization procedure utilizing the general cost function that has been developed in Chapter 6. The GA platform and the D-SPORT controller cost function were fully written using the FORTRAN 90/95 programming language. The written code was developed by taking into account the need for compatibility with the simulation VISSIM-COM interface and the simulation environment, which restricts the use of pre-coded GA tools, such as the GA toolbox in MATLAB. The selection of the FORTRAN 90/95 programming language over other more commonly used languages is based on the ability of FORTRAN 90/95 programs to effectively handle number crunching operations. For the problem at hand, it was assessed that the use of the FORTRAN 90/95 language would provide powerful computational performance and faster computational speeds compared to other programming languages [63]. The FORTRAN 90/95 code that was developed can be found in Appendix B.

Figure 8.1 illustrates a flowchart for the GA process that was implemented. The figure represents the GA optimization platform that appears in Figure 6.3. Once the controller starts the GA optimization process, the GA parameters, such as population size, maximum generations, mutation rate, and crossover rate, are passed to the GA platform. Following reception of these parameters, the GA platform generates the initial population of potential solutions. Traffic and transit data collected from the simulation environment are then fed into the cost function and used to evaluate each individual within the current population of solutions. Following this evaluation, the GA platform then starts the selection process and initiates the generation of the next generation of solutions by apply GA operations, such as mutation, crossover, and elitism, to the current population of solutions.

The above process is iterated until the GA platform reaches the maximum allowable number of generations, or the GA termination criterion is met. Once a near-optimal solution is found, the solution is transmitted to the D-SPORT controller, which implements it in the simulation environment.

The following subsections discuss in more details the modeling and operation of the GA optimization process. Attention is first put on the binary coding and decoding for population strings. This is followed by examinations of the roulette-wheel and tournament selection methods. The GA performance under different mutation and crossover rates is then addressed, and the influence of elitism on the GA operation is discussed.

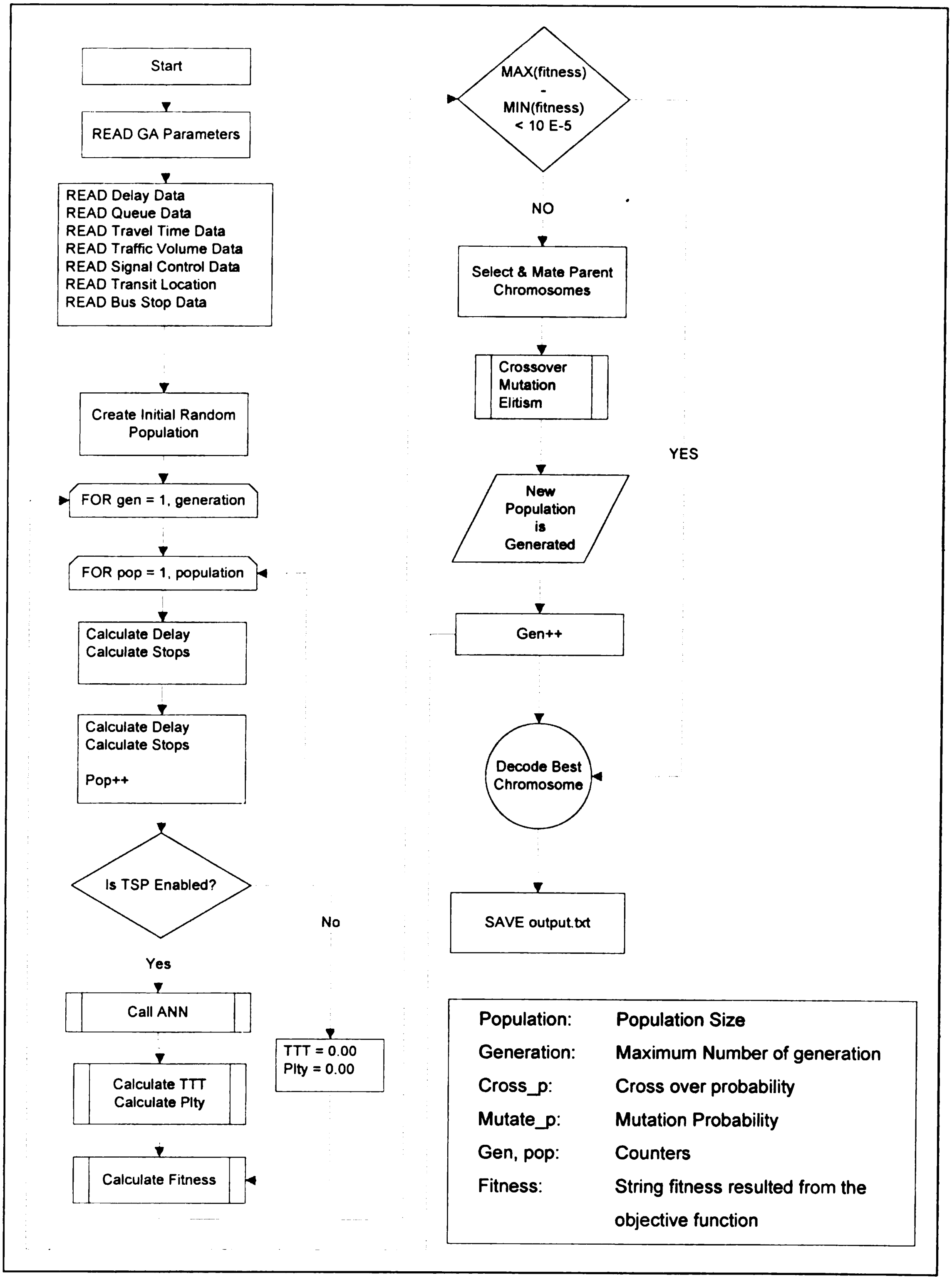


Figure 8.1 The genetic algorithm optimization process flowchart

8.1.1. Solution String Modeling

In GA, each possible solution is represented in a chromosome (also called string). In a classical GA, a binary coding is used to represent the chromosome. Each chromosome consists of a number of bits that could have a value of either 0 or 1. Individual decision variables are represented by a number of bits that is determined based on the potential range of the variable range and its increment. As an example, Figure 8.2 shows a schematic sketch for a chromosome modeling a number of decision variables. In this example, the variables DV_1, DV_2 and DV_3 represent the green time assigned to a specific approach. For the problem at hand, the green time can range between 20.0 and 40.0 seconds, at an increment of 0.5 seconds. Based on this information, 6 bits are found to be sufficient to cover the all the possible values that the variable modeling green time could take.

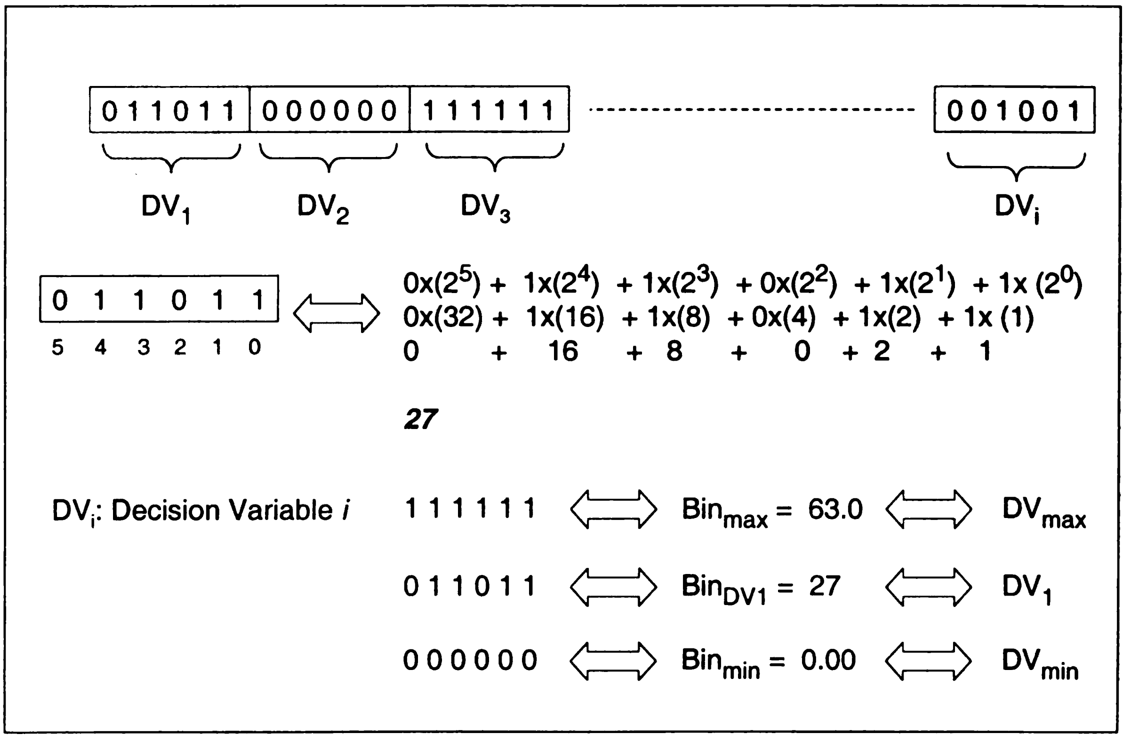


Figure 8.2 Binary system representation and decoding

Conversion from a binary system to a decimal representation is done using the following equation:

$$Bin_{DV_i} = \sum_{p=0}^{P} bin \times 2^{p} \qquad \text{Equation 8.1}$$

Where:

Bin_{DV_i} :	Decimal representation for decision variable i
bin_p :	The bin value in place p
p :	The bin relative location within the chromosome part that represents decision variable i
P :	The total number of bins occupied to represent decision variable i

In Figure 8.2, all the bits associated with the decision variable DV_2 have a value of 0. This corresponds to a decimal representation of a 0.00. For the variable DV_3, the assignment of a value of 1 to all the bits associated to the variable results in a decimal representation of 63.00. The sequence of 0s and 1s associated with the variable DV_1 finally results in a decimal value of 27, as illustrated in the figure.

As defined in Chapter 6, the lowest value a decision variable can take is the minimum allowable green. A value of 0 for the variable DV_2 is therefore out of range as the pre-defined minimum allowable green is 20 seconds. For the variable DV_3, a decimal value of 63.0 is also out of range, being beyond the maximum allowable green time of 40 seconds. To ensure that only valid green times are considered, interpolation is used to determine the actual value of the decision variable value, rounded to the nearest 0.5

second, based on the variable's minimum and maximum value constraints. Equation 8.2 shows the equation used to perform this calculation.

$$DV_i = \frac{INT\left(2 \times \left(\frac{Bin_{DV_i} - Bin_{\min}}{Bin_{\max} - Bin_{\min}} \times \left(DV_{\max} - DV_{\min}\right) + DV_{\min}\right)\right)}{2} \quad \text{... Equation 8.2}$$

Where:

DV_i: Decision variable i actual value

$Bin_{\min}$: The minimum decimal representation a decision variable can ever take, when all the bits are 0's. It has a value of 0.00 with 6 bins per decision variable

$Bin_{\max}$: The maximum decimal representations a decision variable can ever take, when all the bits are 1's. It has a value of 63.00 with 6 bins per decision variable

$DV_{\min}$: The minimum allowable value a decision variable can take, which is equal to the minimum allowable green of 20 seconds

$DV_{\max}$: The maximum allowable value a decision variable can take, which is equal to the maximum allowable green of 40 seconds

8.1.2. Genetic Algorithm Operations

This section details the selected operations for the GA platform that has been developed. Elements described include GA search termination criteria, population size, the selection method, the crossover operator, the mutation operator, the elitism operator, the population size, and the number of generations considered.

For the evaluations, the maximum number of generation was set to 1000 generations, as initial experiments showed that executing longer runs had negligible improvement effects on the cost function values. The population size was further set to 300 chromosomes as initial experiments also showed that a population size of 300 with a maximum of 1000 generations generally provide good progress in the solution search process for the problem at hand.

Figure 8.3 shows the combination of GA operational parameters that were tested prior to selecting a set of preferred parameter values. Parameters considered include two selection methods, five crossover rates, and five mutation rates. Selection of the parameter values to adopt is based on performance of the optimization process on a Dell Dimension 3000 desktop with a central processor unit (CPU) speed of 2.80 GHz and a random access memory (RAM) of 256 MB.

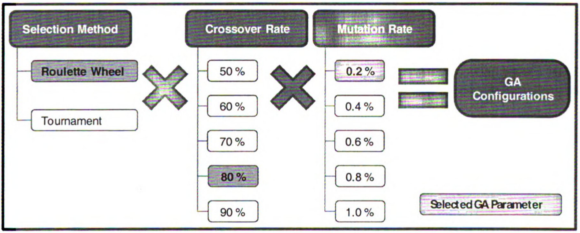


Figure 8.3 The examined genetic algorithms configurations

Termination Criteria

Termination criteria were added to the GA search process to avoid extensive computations yielding little solution improvements. In this case, the GA was instructed to terminate the solution search process if (a) the maximum number of generations is reached or if (b) the absolute difference between the maximum and the minimum cost function value for any generation is found to be less than 10^{-5}.

Selection Method

Selection is the process by which individual solutions (strings) are selected for contribution to the generation of next batch of solutions. In this research, two selection methods are examined, the roulette wheel selection and the tournament selection methods.

The roulette wheel selection algorithm is utilized to perform the parent's string selection process. For each string, the probability of its selection is equal to the ratio of its own cost function value to the sum of cost function values of all strings within the same generation. Once the probability of selection for each string is computed, the cumulative probability is then computed for each string in the generation in the order of their existence within the same generation. The next step involves generating a random number varying between 0.0 and 1.0. This number is used to identify strings with a maximum cumulative cost that is less than or equal to the generated random number. The identified string is then selected and placed in the mating pool. This operation is iterated until all the parents are selected.

For the tournament selection, strings are randomly grouped in pairs, and then from each group, the string with the lowest value for the cost function is chosen for inclusion in the mating pool to generate offspring for the next generation of solutions. For the problem at hand, the binary tournament selection was examined, where each group contains two strings that are selected randomly.

Crossover Rate

Crossover is defined as the genetic process of combining the genes of one chromosome with those of another to create offspring inheriting traits of both parents. For this research, a simple crossover technique is used in which genes are exchanged between two chromosomes around one selected point. For the two selected parents from the mating pool, a random number between 0.0 and 1.0 is first generated. Crossover then only occurs if the generated random number is less than a pre-defined crossover rate.

In this research, different crossover rates are tested for the given population size of 300 and the maximum number of generations of 1000, along with the roulette wheel and tournament selection methods discussed earlier. A total of 5 crossover rates are examined. These rates vary between 50% and 90% with an increment of 10%. Based on the sensitivity analysis of the cost function value to the crossover rate, a value of 80% crossover was selected. The performed sensitivity analysis is discussed in Section 8.1.3.

Mutation Rate

The mutation operation is helpful in a GA search to avoid being trapped within a local optimal solution. Mutation occurs when a randomly selected gene swaps its value, from 0 to 1 or from 1 to 0. For the problem in hand, a bit-wise mutation is used, where the mutation rate identifies the probability of mutated genes, and all the genes in a population have an equal chance of being mutated. In order to determine which bit (gene) will swap its given value, a random number between 0.0 and 1.0 is generated. If the generated random number is less than the mutation rate, then one gene in the population is randomly selected and subjected to the mutation operation. This process is iterated until the expected number of mutated genes is reached. The expected mutated genes can be simply calculated as the total number of bits in the population times the bit-wise mutation rate.

A total of 5 mutation rates were examined. The rates that were examined varied between 0.2% and 1.0%, with an increment of 0.2%. As shown in Figure 8.3, these mutation rates are tested along with the crossover rates and the selection methods that are mentioned earlier in order to determine the GA parameters that will be used to optimize the D-SPORT controller cost function. However, the 0.2% mutation rate is selected as discussed Section 8.1.3.

Elitism

Elitism is defined as keeping the best string of a given generation when developing the next generation of strings. Elitism seeks to retain unchanged one or more strings having the lowest cost function value within a given population to the next population. This is to ensure that the best solution is not lost when going from one generation to the next. For the developed GA platform, elitism is applied based on observations that it results in better GA performance, as discussed in Section 8.1.3.

8.1.3. Genetic Algorithm Operational Performance

Figure 8.4 and Figure 8.5 show the cost function value averaged across 50 runs with different random seeds for different selection methods, crossover and mutation rates, where the objective is to minimize the cost function value. The selected number of runs returns a statistical power of at least 80% for all the pairwise comparisons. These figures show a statistically significant difference between the two selection methods, with a confidence level that exceeds 95%. The results show that the roulette wheel selection methods resulted in lower cost function values compared with the values that resulted from tournament selection. The final GA parameters that are used are the ones associated with the best (lowest) cost function values, which are those associated with the use of roulette wheel selection, with 80% crossover rate, and 0.2% mutation rate.

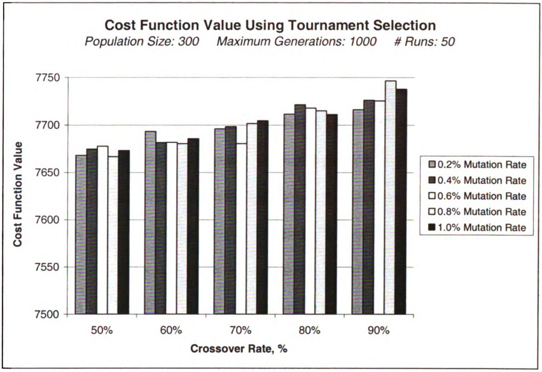


Figure 8.4 Average cost function value for tournament selection, for different crossover and mutation rates

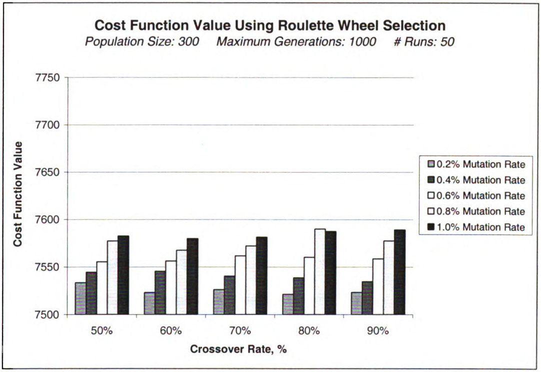


Figure 8.5 Average cost function value for roulette wheel selection, for different crossover and mutation rates

8.1.4. The GA Near-Optimal Solution Search Convergence

The computational performance of the GA platform using the roulette wheel selection is discussed in this section. Figure 8.6 summarizes the total number of function evaluations. For each examined GA configuration, the number of function evaluations is averaged across the 50 GA runs with different random seeds. The number of function evaluations increases with increasing crossover and mutation rates, because the probability of producing and evaluating new individuals is a function of these two rates. As shown in Figure 8.6, the GA configuration with a crossover rate of 90% and a mutation rate of 1% has the highest number of function evaluations, and the configurations with a crossover rate of 50% and a mutation rate of 0.2% has the lowest number of function evaluations.

On the other hand, and as shown in Figure 8.7, there is no clear relationship between the CPU runtime and the number of function evaluations performed. However, it can be seen that with increasing the crossover rate, the runtime tends to be less sensitive to the variation in mutation rate. This is because the probability of evaluating more new individuals increase with higher crossover rate, and applying mutation to these new individuals does not change whether or not they need to be evaluated. In other words, the GA needs to evaluate these new individuals regardless of whether they are mutated or not.

The influence of the crossover rate on the GA search convergence progress versus number of generations and number of function evaluations is examined. The relationship between improvement of the cost function values and the number of generations, and between the cost function values and number of function evaluations performed, are shown in Figure 8.8 and Figure 8.9.

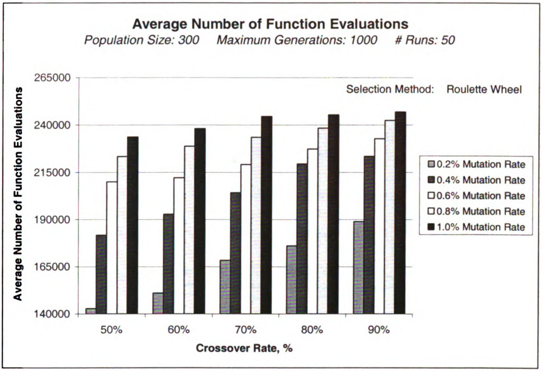


Figure 8.6 Number of function evaluations for different GA configurations

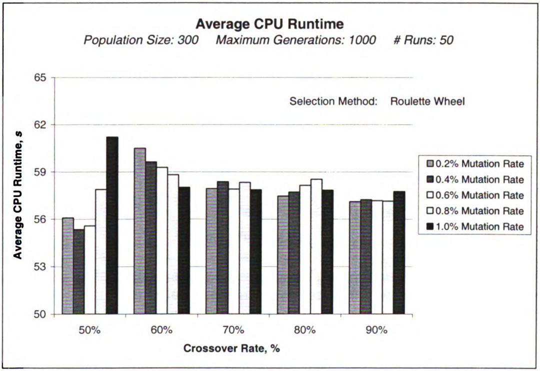


Figure 8.7 Average CPU runtime for different GA configurations

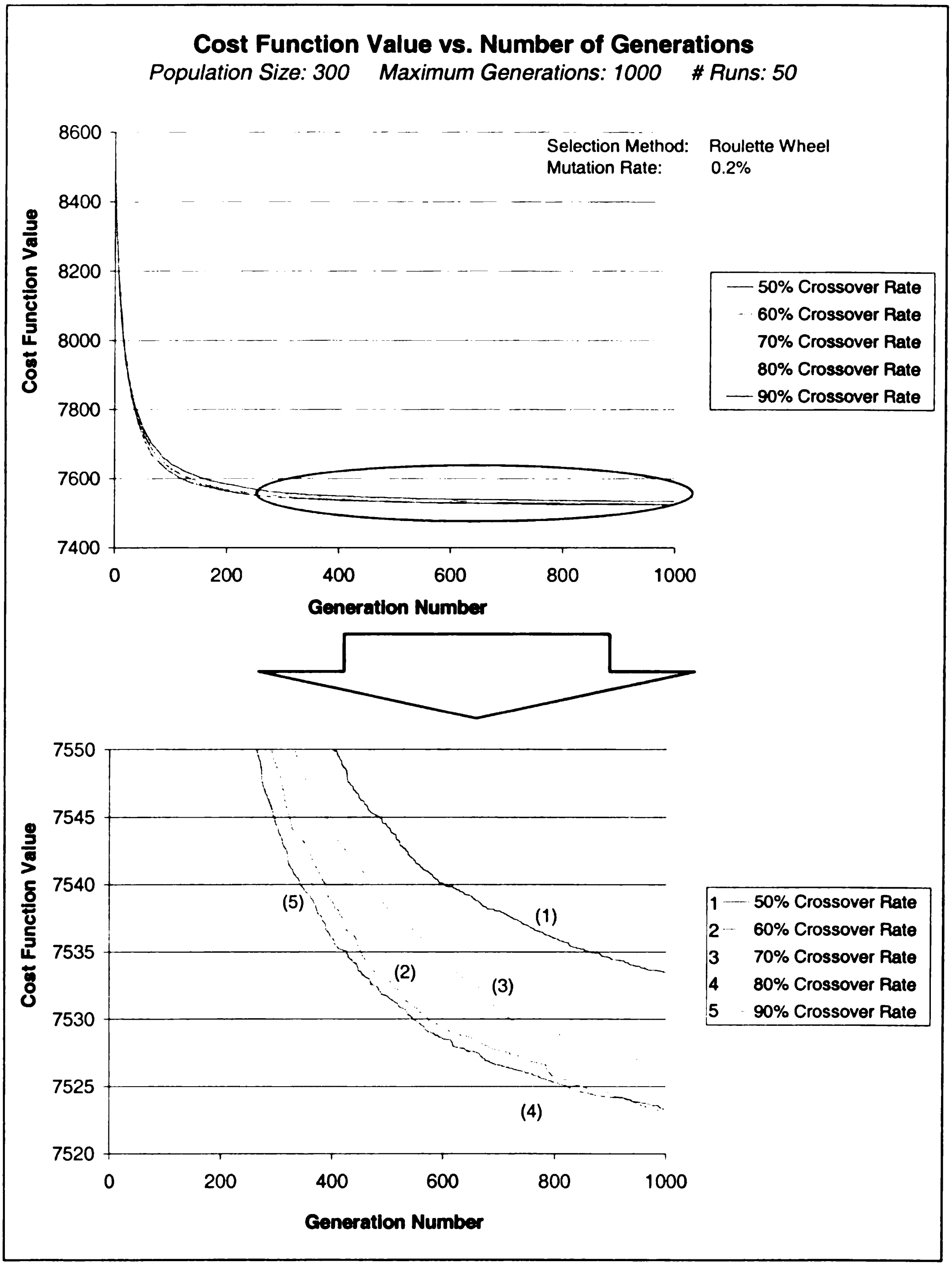


Figure 8.8 GA performance over number of generations for different crossover rates and a mutation rate of 0.2%

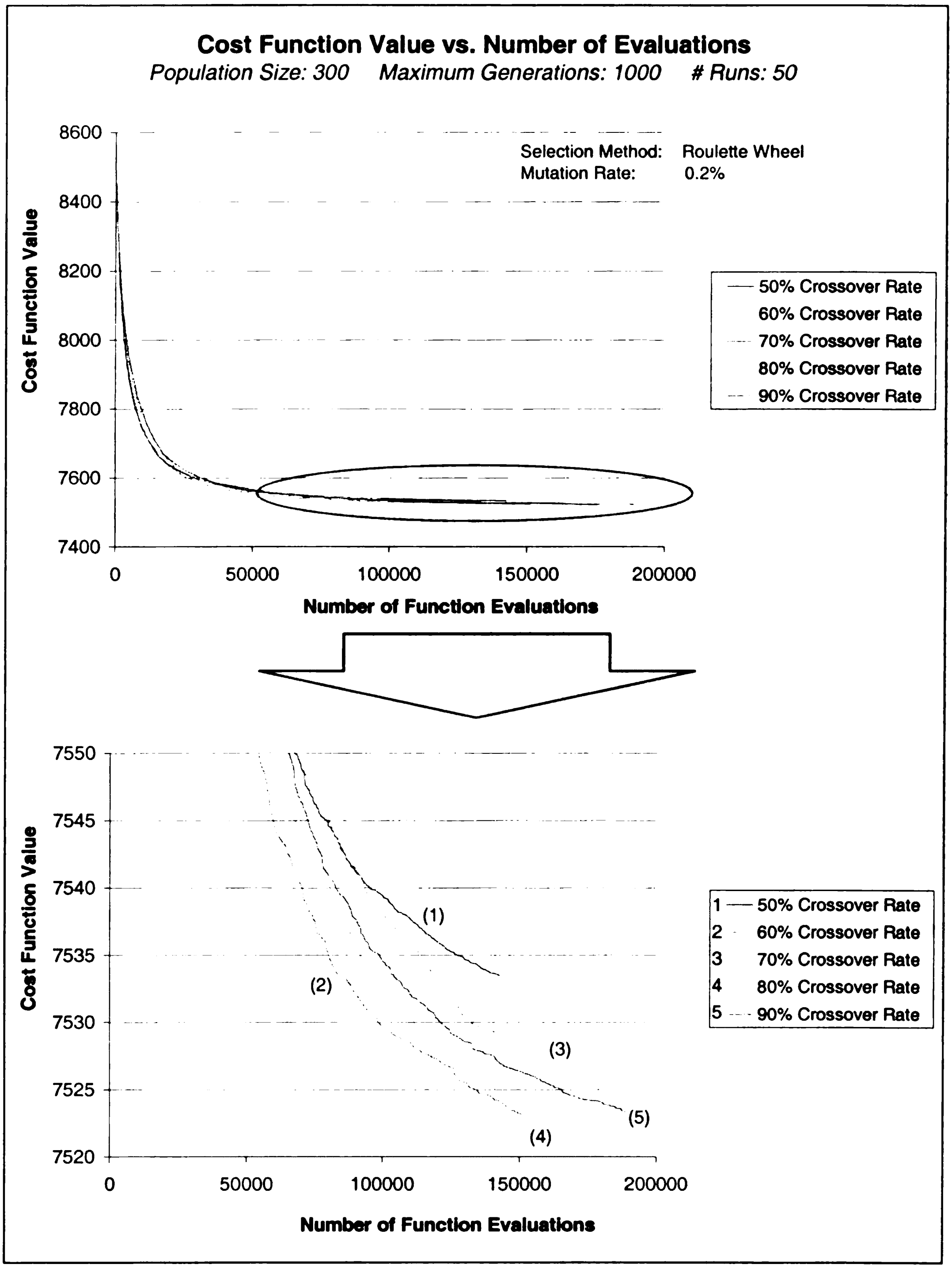


Figure 8.9 **GA performance over number function evaluations performed for different crossover rates and a mutation rate of 0.2%**

Figure 8.8 and Figure 8.9 show that a GA configuration with 80% crossover rate and 0.2% mutation rate converges to a near-optimal solution in early generations. However, and for later generations, there is some improvement observed for that GA configuration.

Figure 8.10 shows the cost function improvement over number of generations with and without applying the elitism operation, when 80% crossover and 0.2% mutation rate are used. The results show that applying elitism has improved the GA convergence.

In conclusion, and according to the previous discussions regarding the performance of the GA under different operational schemes, a GA using roulette wheel selection, 80% crossover rate, 0.2% mutation rate, and elitism was integrated within the simulation environment to perform the required traffic signal timing optimizations.

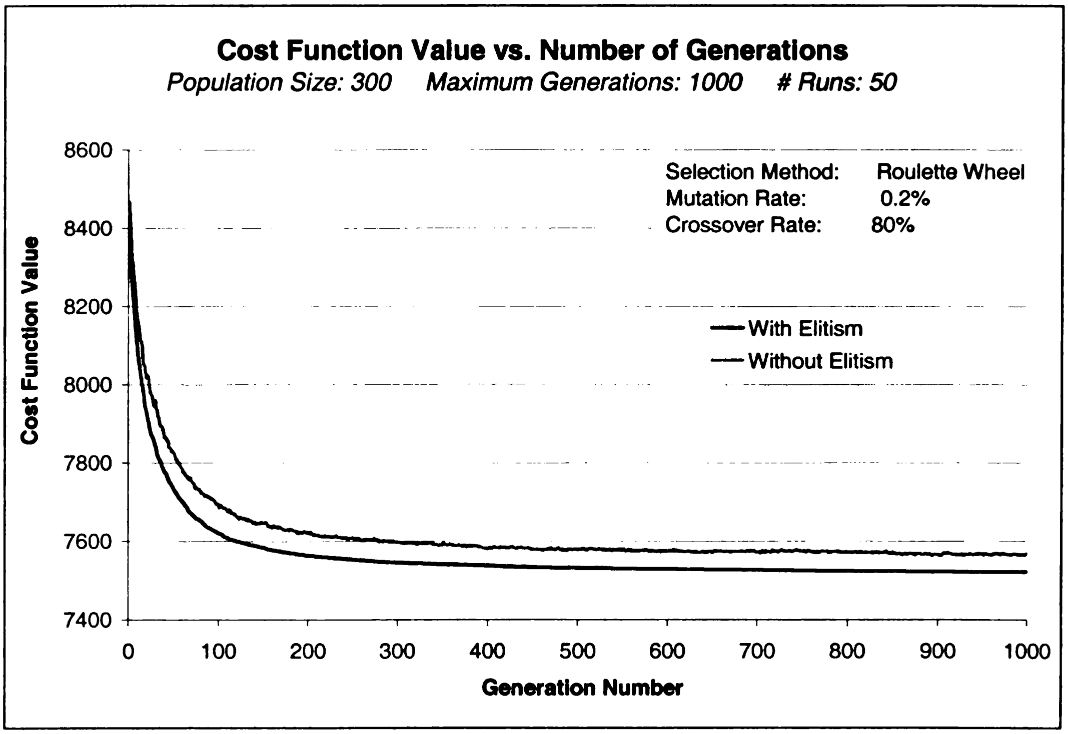


Figure 8.10 GA performance over number of generations with and without Elitism

8.2. Integration of D-SPORT within Simulation Environment

The GA executable file resulted from compiling the FORTRAN 90/95 coding was integrated within the VISSIM microscopic simulation environment. The integration of the GA optimization model was achieved through the software's Component Object Model (VISSIM-COM) interface, which is an external module that enables communication and dynamic object creation between the simulation environment and external processes [58]. This interface facilitates extracting traffic volumes, queuing conditions, bus stop demand, and transit vehicle global position, current signal parameters and local timings, similar to what can be extracted under real-life conditions.

The data required to calculate the cost function value are collected from the simulation environment through the VISSIM-COM interface and then made available to the GA platform. When it is time to call the GA, the VISSIM-COM calls the GA routine, which then reads the *input.txt* file and uses data contained in this file to generate a signal timing solution for the next three signal cycles. Once the signal timings are generated, the VISSIM-COM passes the signal timings to the signal controllers in the simulation environment. The simulation is then continued until it is time to call the GA again.

In order to determine when the GA platform should be called, a countdown counter is used. This counter starts with a time that is equal to the minimum cycle length of the near-optimal solution resulted from a given optimization call. The counter starts counting down till it reaches the zero, in which the GA optimization is called again. This process is iterated throughout the simulation duration. This philosophy is adopted to take into account that the stochastic nature of traffic conditions may create a need for updated solutions in a relatively short term period.

The flowchart shown in Figure 8.11 illustrates the integration process between the D-SPORT model and the simulation environment. The controller collects traffic and transit data, and once it is time to trigger the GA, the controller passes the collected data to the optimizer. The optimizer finds a near-optimal solution based on the passed data, and returns the solution to the controller, which implements the solution within the simulation environment.

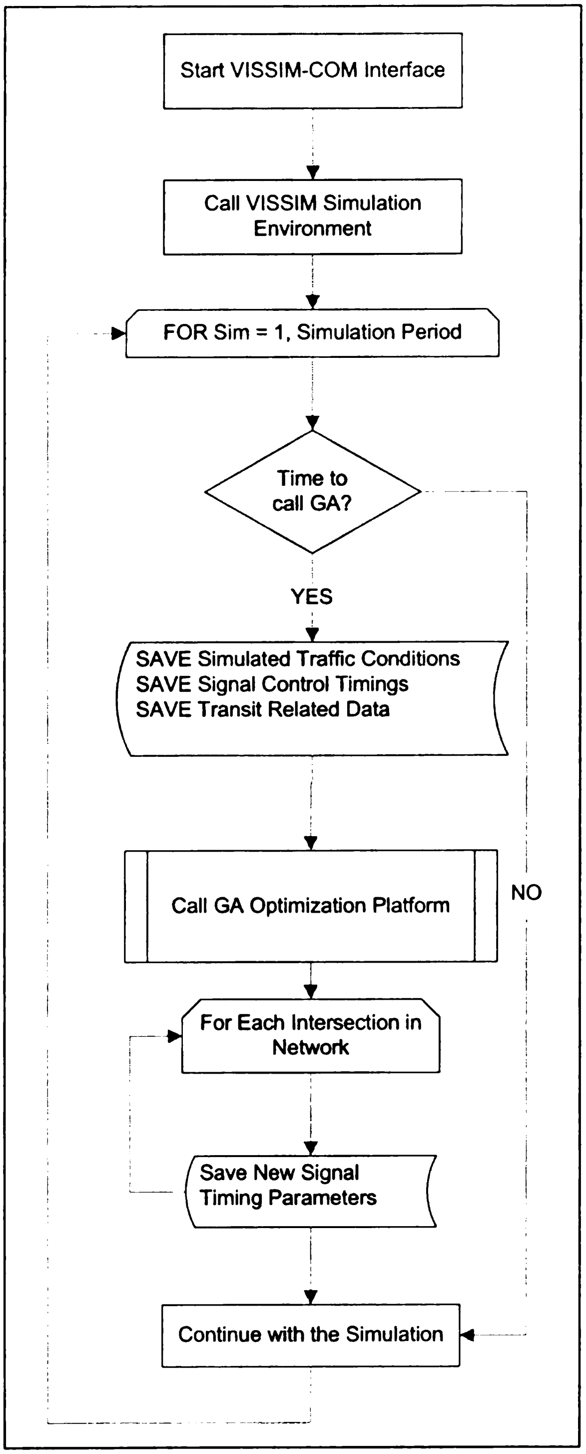


Figure 8.11 The integrated system architecture for the VISSIM-COM Interface within the simulation network framework

8.3. Experimental Setup

This section describes the VISSIM simulation model that was developed to study the performance of the D-SPORT controller. Elements described in this section include the geometric layout for the network, traffic and transit demands, and the alternative traffic control logic considered.

Geometric Layout

The network used to test the model is shown in Figure 8.12. This network consists of 7 intersections and 14 links. It models two intersecting one-way streets with coordinated signals. Link along the two modeled streets features two travel lanes and an exclusive left-turn bay. The Links modeling minor cross-streets have the same features, except that they do not include exclusive turn bays.

Intersections at the upstream ends of the network (intersections 1 and 7) are used to minimize the influence of the boundary conditions and shape arriving traffic into realistic platoons. For this reason, stops and delays incurred at these intersections are excluded from the analyses detailed later in the dissertation. For each approach, detectors are placed at the beginning and the end of the approach to collect traffic volume, delay, and queue length data.

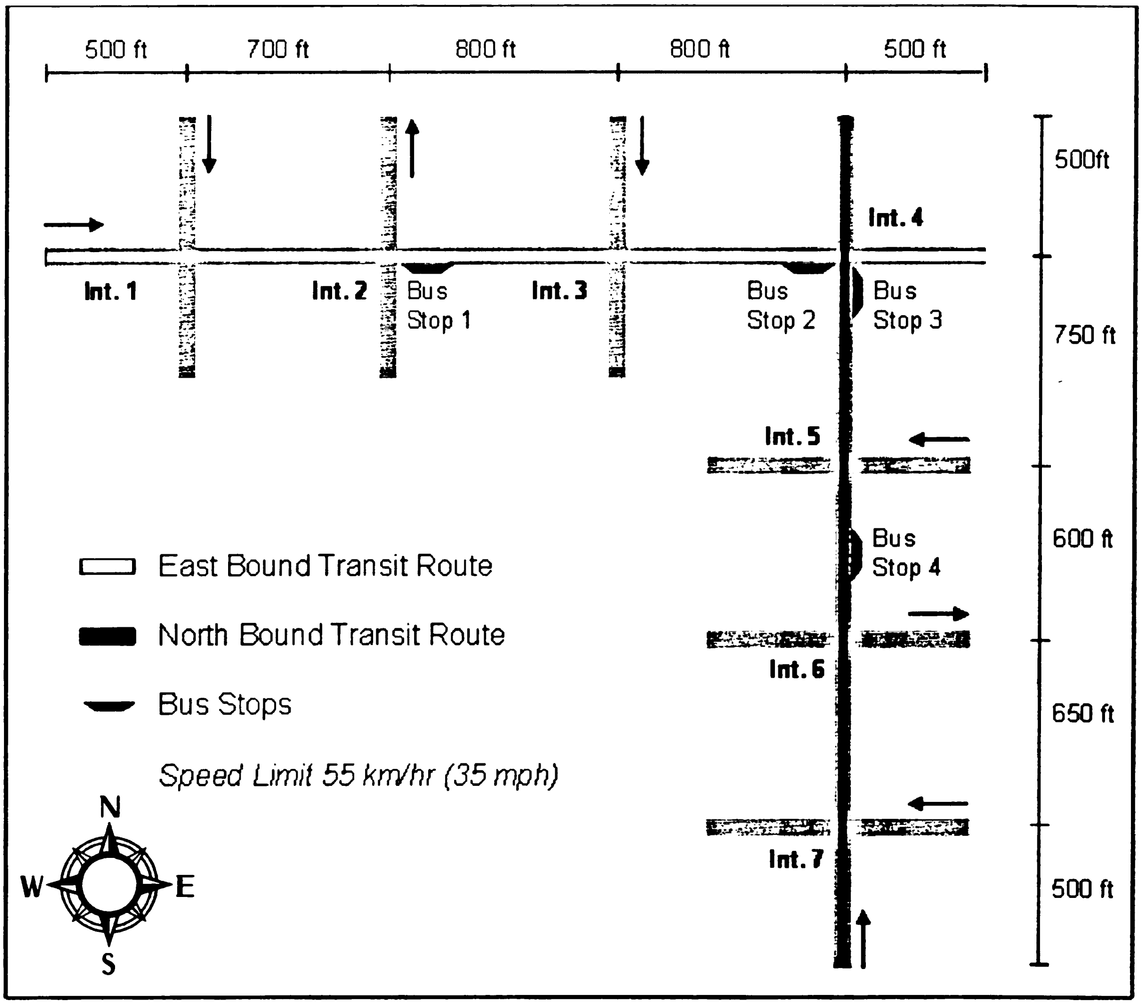


Figure 8.12 Simulated network layout.

Traffic Signal Control

Each intersection is assumed to be controlled by a two-phase signal providing 4-s of amber and a 1-s all red within each phase. The amount of green time assigned to each phase will depend on the type of controller being tested and can be either fixed of variable. Depending on the traffic conditions, the cycle length is allowed to vary between 50 and 90 seconds, and the green time is allowed to vary between 20 and 40 seconds. A 50-second minimum cycle with minimum green restrictions is chosen to ensure that the timing plans being considered do not unduly restrict pedestrian crossing times. Green interval duration for each phase and offsets might vary from cycle to cycle based on the cost function optimization results.

The amount of green time allocated to each phase depends on the type of signal controller being considered. To evaluate the effectiveness of the D-SPORT, scenarios considering various types of controllers were developed:

1- Pre-Timed, which is a pre-timed signal control type hat has an offline optimization signal timing plan with no TSP implementation. This controller is optimized using SYNCHRO 6 for each traffic volume combinations.

2- Pre-Timed TSP, which is a pre-timed signal control, similar to the Pre-Timed controller. However, this controller offers early green and green extension active TSP treatments that are implemented locally. The traffic signal timing plans are optimized using SYNCHRO 6. The standard TSP logic, which only considers average dwell time is the one used to offer green extension and early green active priority treatments, as described in Section 7.6.

3- Real-Time without TSP: this signal control is similar to the proposed GA-Based optimizer while the TSP sub-cost function is inactive. In other words, the parameters α_{PI}, α_{TTT} and α_{SCH} are set to 1.0, 0.0, and 0.0 respectively in Equation 6.6, the equation that represents the D-SPORT controller cost function formula.

4- D-SPORT, the real-time signal control optimization with integrated preferential treatments for transit vehicles.

The first two scenarios represent traditional fixed-operations. They implement signal timings generated before the start of a simulation based on average traffic conditions and that do not change with fluctuations in traffic demand, except for the provision of priority actions to accommodate buses. Signal timings plans in both of these two scenarios were developed using SYNCHRO 6, which is a commonly used traffic signal optimization software.

The scenario against which all other scenarios will be compared is the first one in the list, (i.e., the one featuring fixed-time control without TSP). Comparisons to this scenario will allow to evaluate how buses benefit from the provision of TSP and then to determine the incremental benefits provided by the D-SPORT controller.

Traffic Demand

To remove truck effects, only passenger cars are simulated. Vehicles are generated at the upstream end of each entry link using a Poisson distribution and VISSIM's default vehicle characteristics. To allow testing in a range of traffic conditions, traffic volumes covering low, moderate and heavy traffic conditions along the two major corridors are considered. This meant developing scenarios with volume to capacity ratios (v/c) of 50%, 75%, and 95%. In all scenarios, traffic volumes along the minor streets are kept at a fixed v/c ratio of 50%. The proportion of traffic that turns left or right on each approach is also assumed to be constant, at 15% of the approach volume.

Transit Demand

Two transit routes are assumed traveling along the network, one in the eastbound (EB) direction and one in the northbound (NB) direction. This will allow for the evaluation of the ability of the D-SPORT controller to generate efficient solutions at the intersection of the two one-way streets when conflicting priority needs exist. For each transit route, a transit vehicle is generated to enter the system following a normal distribution with an average headway of 5-minutes and a standard deviation of 20-seconds to take into account the variability associated with the bus entrance to the network. Two bus stops are placed along each transit routes bus routes. Furthermore, it is assumed that transit vehicles are equipped with automatic passenger counts in order to help predict the duration of the dwell time at bus stops, and a Global Positioning System (GPS) to track the location of the transit vehicles while traveling along the network. It is further assumed that all the bus stops have identical boarding rates following a Poisson distribution. Boarding rates considered across various scenarios include rates of 50, 100, and 150 passenger/hr. An alighting rate of 10% of transit vehicle occupancy was finally considered to simulate individuals getting of individual vehicles.

Figure 8.13 summarizes all the possible experiment combinations that were examined. For the total of 72 scenarios, the difference between them is by changing the signal control, the traffic demand along the eastbound and the northbound, and the average passenger demand at bus stops.

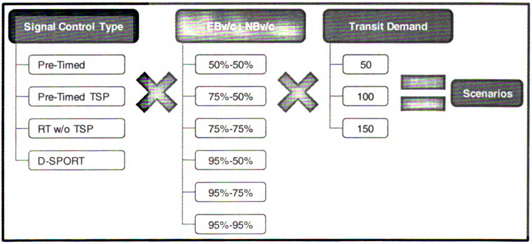


Figure 8.13 Sketch summarizes all the possible combinations that are tested

CHAPTER IX.

SIMULATION RESULTS AND COMPARISONS

In this chapter, the results of the simulation experiments described in Chapter 8 are presented and discussed. The results are broken into different MOEs. Total delay and number of stops are discussed for both traffic and transit movements. Service reliability and schedule adherence are further considered for simulation results pertaining to transit operations. The four signal control algorithms and test network described in described in Section 8.3 are used to evaluate the improvement achieved by the D-SPORT controller over current signal timing optimization practice. To account for stochastic effects in simulations, all the results summarized in this section are based on 10 replications with different random seeds, with each replication covering a 2-hr simulation period.

The simulation results show that the proposed D-SPORT controller is successfully able to respond to the changes in traffic conditions, to provide effective priority to buses while maintaining coordination across adjacent intersections, and to solve difficulties associated with dwell time variability and conflicting priority requests. This success is expressed by reduced transit and general traffic delays and number of stops, as well as by better schedule adherence and service reliability.

9.1. Transit Travel Time Prediction

As described in Section 6.3.2, transit travel time is predicted through macroscopic analysis and ANN predictions as the summation of free flow travel time, dwell time, and delay encountered by a bus along its travel path. To quantify the performance of the developed transit travel time prediction model, the simulation network, traffic volumes and transit demands described in Section 8.3 are used to generate traffic scenarios and measure transit travel times along the two modeled transit routes within each scenario.

While field testing would be an ideal evaluation tested, simulation is used to validate the transit travel time prediction model due to the difficulties in setting up a validation environment in the real world. Conducting evaluations in the real world would require equipping a bus and traffic signal controller with the necessary detection and signal control altering equipment, which is beyond the resources available for this research project.

To ensure the validity of transit travel time predictions, the model was tested with random pre-selected cycle lengths and splits. For each controller, different cycle lengths and splits were randomly generated and assigned for the 2-hour simulation duration. Although the cycle lengths and splits are randomly chosen, they all respect the minimum and maximum allowable green times for individual phases defined in Section 8.3. These parameters were introduced into the VISSIM simulation environment using the VISSIM-COM interface.

While the simulation is running, a snapshot is taken every 60 seconds. When each snapshot is taken, relevant traffic and transit data is collected for each transit vehicle in the network. These data include transit identification number, route, location, speed, data described in Chapter 7 required to predict dwell time activity, and data used to assess travel time volatility. For each transit vehicle in the network when the snapshot is taken, the transit travel time prediction model is used to predict the anticipated time required by the transit vehicle to reach the end of the network. This prediction is then compared with the actual time taken by the vehicle to leave the network to assess performance of the prediction algorithm.

As illustrated in Figure 9.1, predicted transit travel times were found to generally match actual simulated transit travel times. A coefficient of determination (R^2) value of 0.9693 is further computed. Within this context, it is assessed that the model is capable of providing the cost function with a reliable and accurate value for transit travel time prediction model.

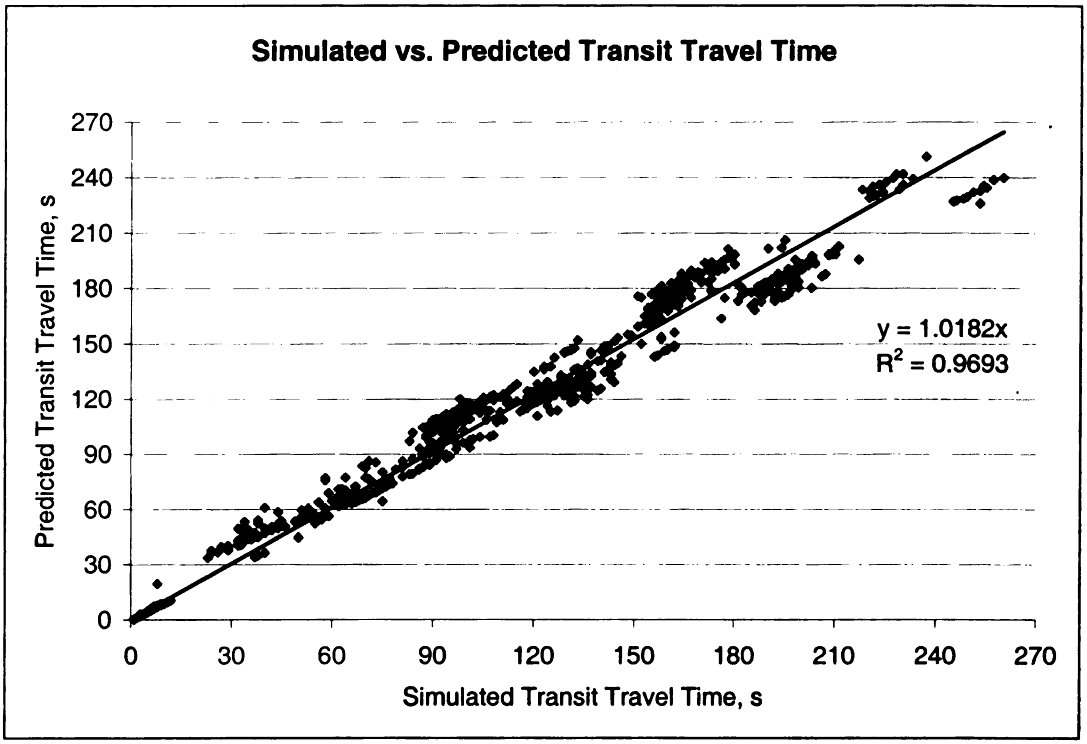


Figure 9.1 Comparison between simulated and predicted Transit Travel Time

9.2. Overall Network Performance

An overall look at network performance is performed by linearly combining average traffic delay, average number of stops per vehicle, average transit travel time, and average transit lateness at bus stops. These parameters are considered based on their inclusion in the D-SPORT controller cost function that is optimized by the GA optimization model (see Equation 6.6) when determining appropriate signal control strategies.

In this research, network performance is estimated by using simulation statistics generated at the end of each 2-hr simulation period. To illustrate how overall network performance is computed, consider the following example. First, it is assumed that the following statistics summarize simulation results from a given scenario:

- Average vehicle delay for each phase group (total of 14 phase groups): 15.0 s/veh
- Average number of stops for each phase group: 1.2 stops/veh
- Average transit travel time for each transit route (total of 2 transit routes): 250.0 s
- Average transit lateness for each bus stop (total of 4 bus stops): 50.0 s

From the statistics described above, the overall network performance can be computed as:

$$COST = \left[\sum PI \quad \sum TTT \quad \sum SCH\right] \cdot \begin{bmatrix} \alpha_{PI} \\ \alpha_{TTT} \\ \alpha_{SCH} \end{bmatrix}$$

$$COST = \left[14 \times 15.0 + 10 \times (14 \times 1.2) \quad 2 \times 250.0 \quad 4 \times 50.0\right] \cdot \begin{bmatrix} 1.0 \\ 10.0 \\ 10.0 \end{bmatrix}$$

Overall Network Cost Function Value = 7378.0

Similarly estimated network cost function values for different levels of transit and traffic demand are shown in Figure 9.2, Figure 9.3, and Figure 9.4. As can be seen in the figures, the proposed D-SPORT controller produced the lowest cost function value in all cases, which is what was desired. This observation is consistent with the GA minimization process, which looks for a solution that minimizes traffic delays, number of stops per vehicle, transit travel time, and transit lateness at bus stops. Although this conclusion is based on an aggregate value of all MOEs considered, the following sections focus more on specific each measure of effectiveness used to calculate the overall network performance.

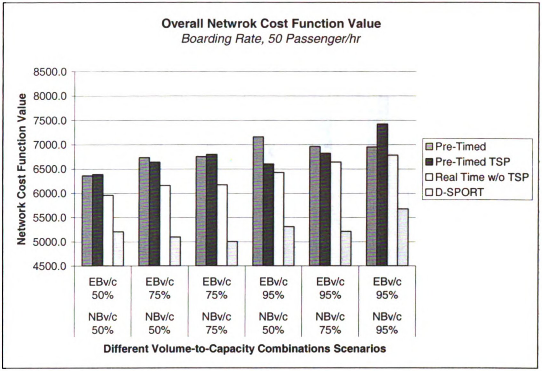


Figure 9.2 Overall traffic network cost function value under different traffic demand and 50 passenger/hr transit demand

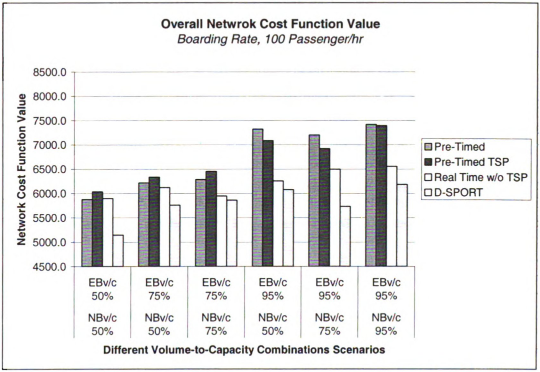


Figure 9.3 Overall traffic network cost function value under different traffic demand and 100 passenger/hr transit demand

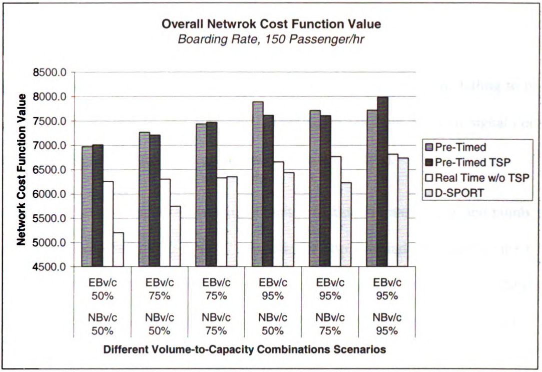


Figure 9.4 **Overall traffic network cost function value under different traffic demand and 150 passenger/hr transit demand**

9.3. Network Total Delay and Number of Stops

The first measures of effectiveness discussed in this section are the total delay and number of stops for transit and non-transit vehicles. Estimates of traffic delays and number of stops are important to emphasize the ability of a system to function efficiently under different traffic conditions, especially during peak-hour. As an example, the level of service (LOS) for a signalized intersection, which is a qualitative measurement of traffic signal operational performance, is measured based on expected incurred delay.

Table 9.1 summarizes transit delays and number of stops averaged across 10 replications under various transit and traffic demand levels and identical signal control alternatives for both the eastbound and northbound routes of the network shown in Figure 8.12, it also shows pair-wise statistical comparisons with scenarios using the pre-timed controller. On the other hand, Figure 9.5, Figure 9.6, and Figure 9.7 visually compare transit delay when different signal controllers are used. Note that the selection of the 10 replications has ensured that the statistical power is at least 80% (i.e., the probability of failing to reject the assumption that delays and number of stops estimates from any two signal control scenarios are the same, when they are actually different is 20%).

In most scenarios, the results show a significant reduction in transit delay and number of stops per vehicle when the D-SPORT controller is implemented compared to the other tested signal control alternatives. While transit delay reduction varies between 20% and 70%, the reduction in number of transit stops varies between 10% and 85%. This finding is associated to the fact that D-SPORT interacts with traffic demands, allowing the D-SPORT controller to adjust in real-time the amount of green time allocated to each signal phase in response to demand fluctuations in a manner that benefits both transit and general traffic. The pre-timed signals, on the other hand, do not react to any fluctuation in traffic demand, with the exception of the provision of early greens and green extensions to accommodate approaching transit vehicles when the TSP logic is active.

Compared to the real-time controller without TSP, the results also show that the D-SPORT controller has resulted in a better transit operational performance in most of the scenarios. These reductions in the amount of delay and number of stops per vehicle are associated with a minimization of the negative influence of transit vehicles on the non-transit vehicles by providing the traffic stream with better signal timings and better traffic progression.

In some scenarios, the implementation of the D-SPORT controller resulted in higher delay for one transit route compared to the other scenarios, especially in cases in which there is a high difference in traffic congestion between the eastbound and northbound directions of the simulated network. This is because the D-SPORT controller attempts to improve overall system performance by acting on a set of sub-cost benefits. For instance, the D-SPORT controller may not find an ideal timing plan to reduce transit delay, but could develop a plan that significantly improve non-transit traffic or improve schedule adherence by imposing a short additional delay to the transit vehicles. By considering network–level traffic, taking into account all the transit routes in the network, and all the sub-objectives considered, such as service reliability and schedule adherence, it can be easily noticed that the D-SPORT controller model performs better than the alternative signal control approaches considered.

As an example, consider the scenario with an eastbound v/c ratio of 95%, and northbound v/c ratio of 50%. In this case, higher eastbound transit delays are observed compared to other scenarios. However, the average transit delay obtained by considering the two transit routes is lower than other scenarios. This can also be observed within the context of the cost function evaluation, where the controller that produced the lowest overall network cost is the D-SPORT controller. For instance, the cost function value for the D-SPORT controller is 5206.04, while cost function values of 6361.31, 6389.70, and 5959.64 are associated with the pre-timed, pre-timed with TSP, and real-time without TSP, respectively, for the transit demand level of 50 passenger/hr.

Table 9.1 Simulated transit delays and number of stops for different traffic signal controls

Boarding Rate Passenger/hr		PreTimed (1)		Pretimed TSP (2)		Rea-Time w/o TSP (3)		D-SPORT (4)	
		Delay s/veh	# Stops stops/veh	Delay s/veh	# Stops stops/veh	Delay s/veh	# Stops stops/veh	Delay s/veh	# Stops stops/veh
v/c EB 50% - v/c NB 50%									
50	EB Transit	70.56	2.52	71.95	2.55	51.66 *	1.33 *	45.80 *	1.13 *
50	NB Transit	18.26	0.89	18.16	0.89	7.81 *	0.26 *	7.20 *	0.25 *
100	EB Transit	22.29	0.88	29.53 *	1.58 *	31.20 *	1.92 *	26.70 *	1.70 *
100	NB Transit	8.96	0.36	8.27 *	0.32 *	7.80 *	0.26 *	7.63 *	0.30 *
150	EB Transit	66.59	2.34	66.38	2.33	25.00 *	1.00 *	11.60 *	0.10 *
150	NB Transit	17.45	0.85	17.38	0.84	7.10 *	0.22 *	5.44 *	0.25 *
v/c EB 75% - v/c NB 50%									
50	EB Transit	73.62	2.68	73.5	2.54 *	51.94 *	1.41 *	48.50 *	1.29 *
50	NB Transit	20.99	1.05	21.79 ^	1.08	13.27 *	0.49 *	10.30 *	0.34 *
100	EB Transit	30.69	1.43	39.48 *	1.78 *	36.50 *	2.00 *	27.10 *	1.70 *
100	NB Transit	9.33	0.38	10.83 *	0.52 *	12.40 *	0.42 *	9.35	0.38
150	EB Transit	68.31	2.44	68.53	2.46	26.30 *	1.00 *	13.40 *	0.40 *
150	NB Transit	19.79	0.98	19.99	1.00	11.40 *	0.36 *	5.97 *	0.30 *
v/c EB 75% - v/c NB 75%									
50	EB Transit	72.68	2.64	83.97 *	2.69	53.22 *	1.45 *	51.70 *	1.42 *
50	NB Transit	21.23	1.04	21.57	1.13 *	13.05 *	0.47 *	11.40*	0.42 *
100	EB Transit	33.51	1.55	39.33 *	1.80 *	26.70 *	2.08 *	26.40 *	1.70 *
100	NB Transit	9.71	0.38	10.71 *	0.49 *	12.60 *	0.48 *	6.88 *	0.36
150	EB Transit	67.83	2.43	68.09	2.46	28.80 *	1.21 *	14.00 *	0.50 *
150	NB Transit	20.15	0.98	20.30	0.99	13.10 *	0.44 *	6.60 *	0.36 *
v/c EB 95% - v/c NB 50%									
50	EB Transit	113.37	2.59	79.80 *	1.79 *	49.93 *	1.34 *	53.20 *	1.38 *
50	NB Transit	16.62	1.03	18.33 *	1.16 *	18.20 *	0.71 *	16.70	0.66 *
100	EB Transit	79.64	2.17	60.00 *	1.20 *	33.40 *	2.00 *	27.60 *	1.40 *
100	NB Transit	9.84	0.44	11.61 *	0.54 *	21.30 *	0.99 *	14.92 *	0.72 *
150	EB Transit	107.73	2.57	106.76	2.53	24.80 *	0.79 *	14.50 *	0.10 *
150	NB Transit	16.09	0.99	16.17	0.99	18.4 *	0.71 *	12.65 *	0.65 *
v/c EB 95% - v/c NB 75%									
50	EB Transit	66.35	2.52	60.44 *	2.47	53.54 *	1.53 *	49.00 *	1.54 *
50	NB Transit	20.35	1.29	21.09 ^	1.36	18.59 *	0.74 *	15.90 *	0.58 *
100	EB Transit	73.25	2.11	45.36 *	1.19 *	35.20 *	2.33 *	23.40 *	1.60 *
100	NB Transit	15.15	0.78	17.01 *	0.92 *	18.80 *	0.84 *	13.31 *	0.76
150	EB Transit	64.45	2.46	64.02	2.43	25.20 *	1.04 *	12.70 *	0.40 *
150	NB Transit	19.62	1.23	19.81	1.25	18.50	0.72 *	10.74 *	0.54 *
v/c EB 95% - v/c NB 95%									
50	EB Transit	50.05	2.48	63.14 *	3.33 ^	61.77 *	1.77 *	59.00 *	1.67 *
50	NB Transit	27.28	1.76	32.43 *	2.13 *	18.98 *	0.74 *	16.90 *	0.64 *
100	EB Transit	52.19	1.76	59.26 *	3.05 *	29.20 *	1.75	30.90 *	1.70
100	NB Transit	28.52	1.92	37.11 *	2.93 *	19.60 *	0.75 *	16.05 *	0.70 *
150	EB Transit	49.2	2.37	51.3	2.68	27.40 *	1.17 *	15.00 *	0.20 *
150	NB Transit	26.97	1.75	27.82	1.82	17.30 *	0.63 *	9.72 *	0.44 *

* Statistically significant at 99% confidence level compared to case (1), pre-timed with no active TSP

^ Statistically significant at 95% confidence level compared to case (1), pre-timed with no active TSP

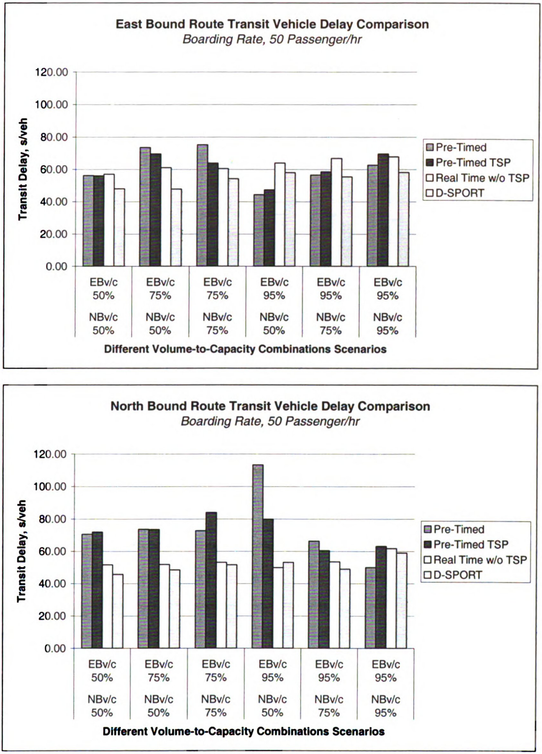


Figure 9.5 Simulated transit delay for different volume-to-capacity combinations and a boarding rate of 50 passenger/hr

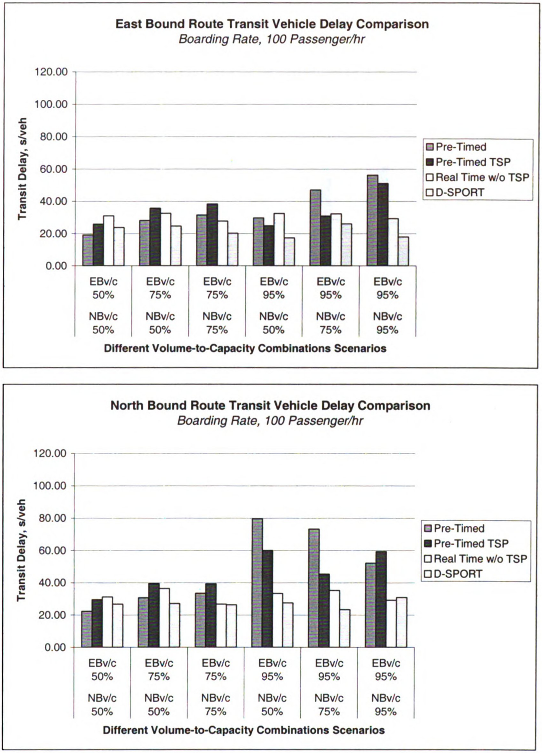
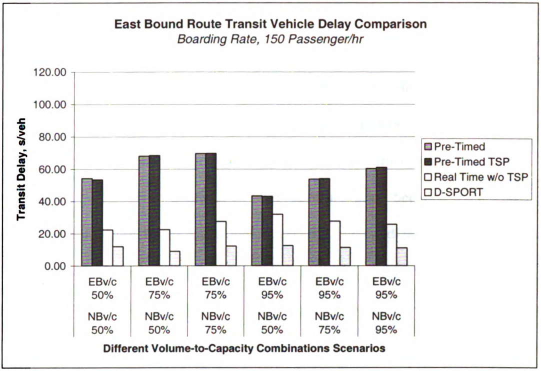


Figure 9.6 Simulated transit delay for different volume-to-capacity combinations and a boarding rate of 100 passenger/hr



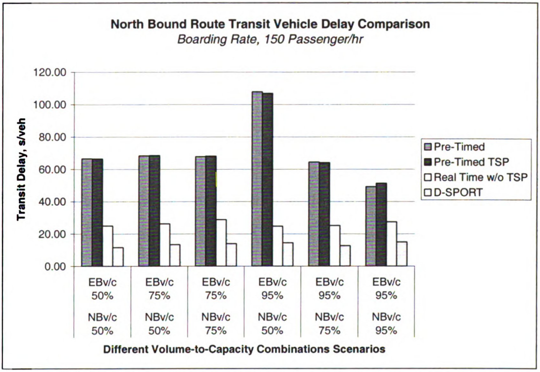


Figure 9.7 Simulated transit delay for different volume-to-capacity combinations and a boarding rate of 150 passenger/hr

While performance evaluation results for non-transit traffic are summarized in Table 9.2, through Table 9.4, Figure 9.8 through Figure 9.13 show the trend of these selected MOEs for different traffic and transit demand levels. These results show that the D-SPORT controller is able to significantly reduce traffic delay and number of stops along the two major arterials compared to the other signal control scenarios considered. While the reductions in traffic delay under D-SPORT control varies between 20% and 80%, reduction in number of traffic stops varies between 10% and 80%. The reductions in traffic delay and number of stops are attributed to the enhanced ability of the D-SPORT controller to respond to traffic fluctuation in real-time and better tailor the signal timing to the current traffic and transit needs. It is also observed that the D-SPORT controller impact on traffic delay and number of stops on the streets crossing the two main corridors varies between non-noticeable to slightly negative impacts.

In cases in which the pre-timed controller is used, the simulation results show a lower delay for the busiest direction when the v/c ratio difference for the eastbound and northbound major directions is wide. For instance, when the eastbound v/c ratio is 0.95, and the northbound v/c is 0.5, the eastbound transit delay varies between 10-17 seconds, while the northbound transit delay varies between 20-35 seconds. Although these results seem to be out of order, they can be explained if the optimization analogy in coordinated network is clarified. As indicated in Section 8.3, the pre-timed signal timings where obtained by using SYNCHRO 6. In this model, the optimization process generally favors the direction with the higher volume by providing this direction with a wider green window, whether the low-volume approach is progressed or not. Therefore, this often results in poor coordination along the low-volume approach.

In the above example, the signal timings produced by SYNCHRO 6 resulted in more green time being allocated to the eastbound approach at intersection 4 than the northbound approach since the former has a higher volume. A problem then occurs when the northbound traffic at intersection 5 along the northbound corridor is being allocated more green time than the northbound traffic at intersection 4, because the eastbound volume is higher than the northbound volume. When the northbound green time at intersection 5 is fully utilized, the signal at intersection 4 will then be unable to process all the vehicles arriving at the intersection within a cycle, causing additional delay.

The D-SPORT controller shows a superior performance compared to the other controllers through its ability to react to traffic demand fluctuation between the two major approaches, which can be translated into a reduction of both transit and general traffic delay and number of stops. In this case, the simulation results show that the approach with higher traffic volume is associated with higher transit delay, which is a logical result. This is because of the better consideration of traffic progression throughout the network and the ability to reduce the disturbance caused by transit vehicles to the traffic stream, which is achieved by the near-optimal green times allocation throughout the network that takes into account both transit and non-transit vehicles movements.

Because of the system-based optimization approach, the D-SPORT controller does not necessarily penalizes traffic on the crossing streets when providing priority treatments to transit vehicles along the two main corridors. The controller attempts instead to synchronize the green timing with the transit expected arrival at bus stops. If green times synchronizing is not possible, more un-utilized green times could be observed, which could potentially lead to unnecessary traffic delay and number of stops.

Table 9.2 **Simulated traffic delay and number of stops for boarding rate of 50 passenger/hr**

	Pre- Timed (1)		Pre- Timed TSP (2)		Rea-Time w/o TSP (3)		D-SPORT (4)	
	Delay *s/veh*	# Stops *stop/veh*	Delay *s/veh*	# Stops *stop/veh*	Delay *s/veh*	# Stops *stop/veh*	Delay *s/veh*	# Stops *stop/veh*
				$EB_{v/c}$ 50% - $NB_{v/c}$ 50%				
EB Traffic	18.26	0.89	18.16	0.89	7.81 *	0.26 *	7.20 *	0.25 *
NB Traffic	14.25	0.7 0	14.82 *	0.71	5.32 *	0.16 *	5.00 *	0.16 *
Crossing EB	9.77	0.54	9.91	0.54	11.98 *	0.55 ^	11.33 *	0.52
Crossing NB	9.74	0.54	9.77	0.54	11.58 *	0.55	11.03 *	0.53
				$EB_{v/c}$ 75% - $NB_{v/c}$ 50%				
EB Traffic	20.99	1.05	21.79 ^	1.08	13.27 *	0.49 *	10.30 *	0.34 *
NB Traffic	17.68	0.79	16.9 *	0.76 *	5.42 *	0.16 *	5.00 *	0.15 *
Crossing EB	12.32	0.6	12.3	0.6	12.23	0.56 *	11.66 *	0.54 *
Crossing NB	10.08	0.54	10.35	0.55	11.53 *	0.55	11.1 *	0.54
				$EB_{v/c}$ 75% - $NB_{v/c}$ 75%				
EB Traffic	21.23	1.04	21.57	1.13 *	13.05 *	0.47 *	11.4 *	0.42 *
NB Traffic	17.92	0.84	29.71 *	1.42 *	9.32 *	0.29 *	8.00 *	0.23 *
Crossing EB	12.31	0.60	12.42	0.60	12.09	0.56 *	11.51 *	0.54 *
Crossing NB	12.1 0	0.60	12.17	0.60	11.60 *	0.55 *	11.10 *	0.53 *
				$EB_{v/c}$ 95% - $NB_{v/c}$ 50%				
EB Traffic	16.62	1.03	18.33 *	1.16 *	18.20 *	0.71 *	16.70	0.66 *
NB Traffic	35.61	1.04	39.46 *	1.20 *	5.47 *	0.16 *	5.00 *	0.14 *
Crossing EB	25.31	0.82	25.35	0.82	12.28 *	0.56 *	11.73 *	0.54 *
Crossing NB	12.78	0.54	14.09 *	0.58 *	11.56 *	0.55	11.14 *	0.53
				$EB_{v/c}$ 95% - $NB_{v/c}$ 75%				
EB Traffic	20.35	1.29	21.09 ^	1.36	18.59 *	0.74 *	15.9 *	0.58 *
NB Traffic	26.20	1.05	27.61 *	1.16 *	8.94 *	0.27 *	8.10 *	0.24 *
Crossing EB	22.60	0.78	22.69	0.79	12.32 *	0.57 *	11.71 *	0.55 *
Crossing NB	17.70	0.68	18.17	0.69	11.77 *	0.56 *	11.16 *	0.54 *
				$EB_{v/c}$ 95% - $NB_{v/c}$ 95%				
EB Traffic	27.28	1.76	32.43 *	2.13 *	18.98 *	0.74 *	16.9 *	0.64 *
NB Traffic	28.01	1.95	43.9 *	3.51 *	13.9 *	0.48 *	12.7 *	0.41 *
Crossing EB	22.69	0.79	22.70	0.79	12.14 *	0.56 *	11.75 *	0.55 *
Crossing NB	22.39	0.78	22.79	0.79	11.73 *	0.56 *	11.1 *	0.54 *

* Statistically significant at 99% confidence level compared to case (1), pre-timed with no active TSP
^ Statistically significant at 95% confidence level compared to case (1), pre-timed with no active TSP

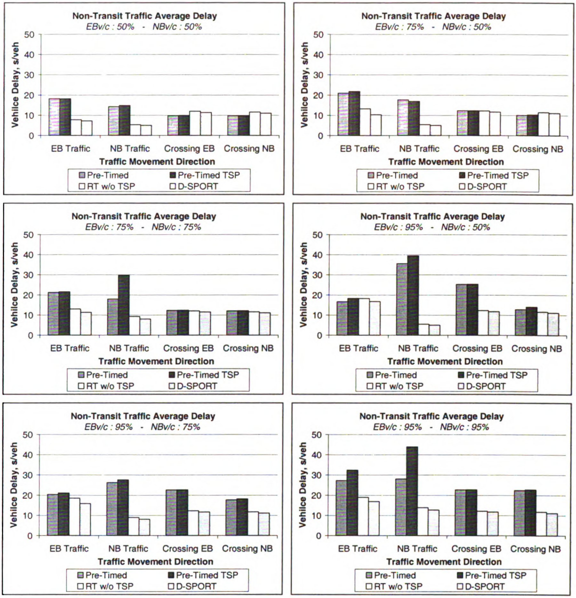


Figure 9.8 Simulated non-transit traffic delay for different v/c combinations and a boarding rate of 50 passenger/hr

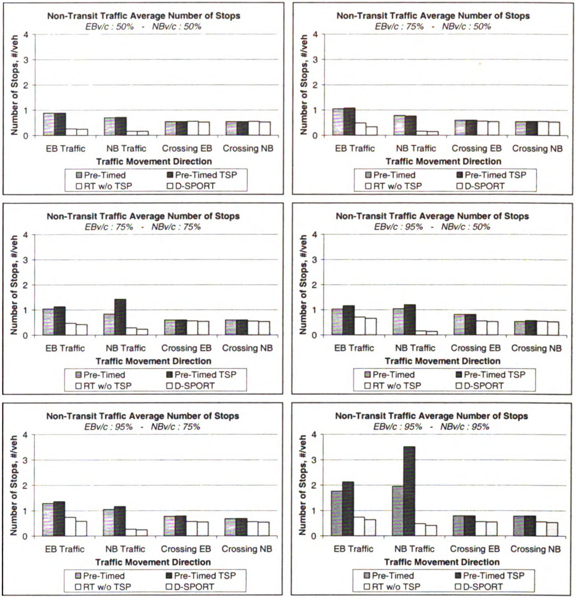


Figure 9.9 Simulated non-transit number of stops for different v/c combinations and a boarding rate of 50 passenger/hr

Table 9.3 Simulated traffic delay and number of stops for boarding rate of 100 passenger/hr

	Pre- Timed (1)		Pre- Timed TSP (2)		Rea-Time w/o TSP (3)		D-SPORT (4)	
	Delay	# Stops	Delay	# Stops	Delay	# Stops	Delay	# Stops
	s/veh	*stop/veh*	*s/veh*	*stop/veh*	*s/veh*	*stop/veh*	*s/veh*	*stop/veh*
				$EB_{v/c}$ 50% - $NB_{v/c}$ 50%				
EB Traffic	8.96	0.36	8.27 *	0.32 *	7.80 *	0.26 *	7.63 *	0.30 *
NB Traffic	4.07	0.16	5.27 *	0.22 *	4.90 *	0.16	4.97 *	0.16
Crossing EB	9.84	0.54	10.09	0.55	11.94 *	0.55	10.84 *	0.55
Crossing NB	9.70	0.54	9.90	0.55	11.86 *	0.56 *	11.01 *	0.54
				$EB_{v/c}$ 75% - $NB_{v/c}$ 50%				
EB Traffic	9.33	0.38	10.83 *	0.52 *	12.40 *	0.42 *	9.35	0.38
NB Traffic	6.65	0.23	8.32 *	0.35 *	5.50 *	0.17 *	5.34 *	0.19 *
Crossing EB	12.34	0.60	12.59	0.61	12.36	0.58 *	11.79 *	0.55 *
Crossing NB	10.06	0.54	10.81 *	0.55 ^	11.43 *	0.54	10.77 *	0.53
				$EB_{v/c}$ 75% - $NB_{v/c}$ 75%				
EB Traffic	9.71	0.38	10.71 *	0.49 *	12.6 *	0.48 *	6.88 *	0.36
NB Traffic	6.05	0.18	9.21 *	0.39 *	11.00 *	0.37 *	10.01 *	0.33 *
Crossing EB	12.32	0.60	12.67 ^	0.61	12.2	0.57 *	11.43 *	0.55 *
Crossing NB	12.08	0.60	12.51 *	0.61	11.59 *	0.55 *	10.85 *	0.55 *
				$EB_{v/c}$ 95% - $NB_{v/c}$ 50%				
EB Traffic	9.84	0.44	11.61 *	0.54 *	21.3 *	0.99 *	14.92 *	0.72 *
NB Traffic	20.04	0.46	27.61 *	0.69 *	5.10 *	0.14 *	5.52 *	0.20 *
Crossing EB	25.35	0.82	25.78	0.82	12.01 *	0.56 *	10.53 *	0.54 *
Crossing NB	12.73	0.55	15.42 *	0.6 *	11.67 *	0.55	11.04 *	0.55
				$EB_{v/c}$ 95% - $NB_{v/c}$ 75%				
EB Traffic	15.15	0.78	17.01 *	0.92 *	18.8 *	0.84 *	13.31 *	0.76
NB Traffic	16.88	0.57	18.80 *	0.74 *	8.7 *	0.27 *	10.24 *	0.36 *
Crossing EB	22.67	0.78	28.60	1.18	12.44 *	0.58 *	10.46 *	0.56 *
Crossing NB	17.70	0.68	24.58	1.11	11.59 *	0.55 *	10.85 *	0.54 *
				$EB_{v/c}$ 95% - $NB_{v/c}$ 95%				
EB Traffic	28.52	1.92	37.11 *	2.93 *	19.6 *	0.75 *	16.05 *	0.70 *
NB Traffic	20.61	1.30	39.38 *	3.1 *	12.2 *	0.39 *	15.27 *	0.62 *
Crossing EB	22.74	0.79	35.71	1.69	11.54 *	0.54 *	10.45 *	0.53 *
Crossing NB	22.44	0.78	35.88	1.68	11.5 *	0.54 *	10.82 *	0.54 *

* Statistically significant at 99% confidence level compared to case (1), pre-timed with no active TSP
^ Statistically significant at 95% confidence level compared to case (1), pre-timed with no active TSP

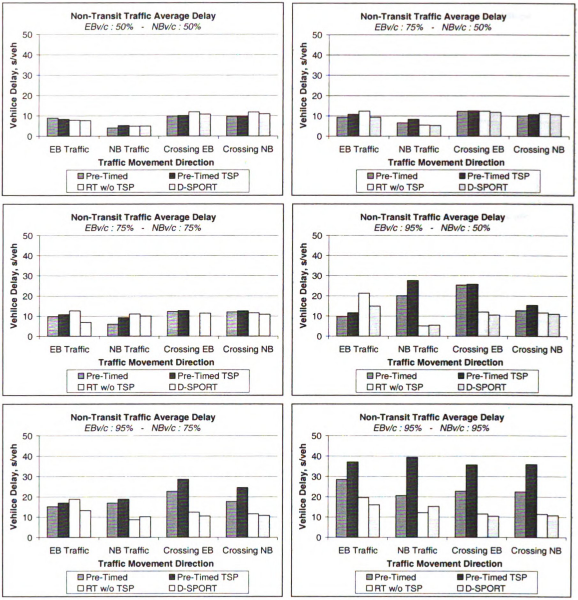


Figure 9.10 Simulated non-transit traffic delay for different v/c combinations and a boarding rate of 100 passenger/hr

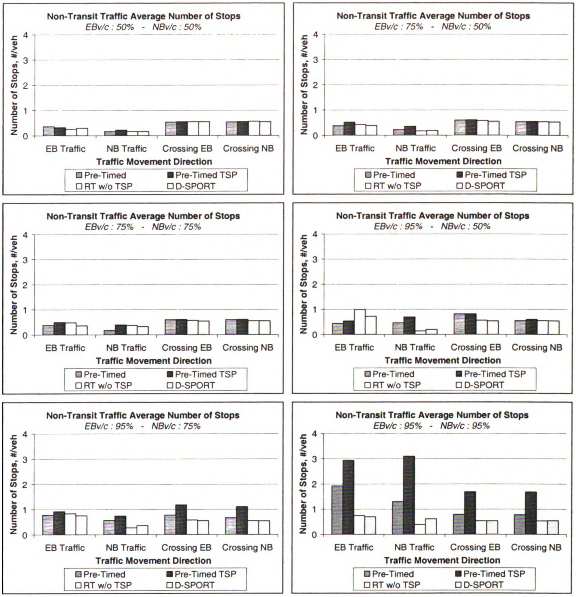


Figure 9.11 Simulated non-transit number of stops for different v/c combinations and a boarding rate of 100 passenger/hr

Table 9.4 **Simulated traffic delay and number of stops for boarding rate of 150 passenger/hr**

	Pre- Timed (1)		Pre- Timed TSP (2)		Rea-Time w/o TSP (3)		D-SPORT (4)	
	Delay s/veh	# Stops stop/veh	Delay s/veh	# Stops stop/veh	Delay s/veh	# Stops stop/veh	Delay s/veh	# Stops stop/veh
				$EB_{v/c}$ 50% - $NB_{v/c}$ 50%				
EB Traffic	17.45	0.85	17.38	0.84	7.10 *	0.22 *	5.44 *	0.25 *
NB Traffic	13.21	0.65	13.30	0.65	5.40 *	0.17 *	4.86 *	0.16 *
Crossing EB	9.81	0.54	9.82	0.54	12.3 *	0.56 *	11.37 *	0.54
Crossing NB	9.74	0.54	9.74	0.54	11.36 *	0.53 ^	11.26 *	0.55
				$EB_{v/c}$ 75% - $NB_{v/c}$ 50%				
EB Traffic	19.79	0.98	19.99	1.00	11.4 *	0.36 *	5.97 *	0.30 *
NB Traffic	16.50	0.73	16.70	0.75	5.80 *	0.17 *	5.16 *	0.17 *
Crossing EB	12.31	0.60	12.33	0.60	11.80 *	0.54 *	12.08	0.57 *
Crossing NB	10.08	0.54	10.13	0.54	11.43 *	0.54	11.10 *	0.56 *
				$EB_{v/c}$ 75% - $NB_{v/c}$ 75%				
EB Traffic	20.15	0.98	20.3	0.99	13.1 *	0.44 *	6.60 *	0.36 *
NB Traffic	16.84	0.78	17.11	0.79	10.3 *	0.34 *	10.00 *	0.32 *
Crossing EB	12.31	0.60	12.32	0.60	11.79 *	0.55 *	9.96 *	0.52 *
Crossing NB	12.12	0.60	12.14	0.60	11.69 *	0.55 *	11.12 *	0.55 *
				$EB_{v/c}$ 95% - $NB_{v/c}$ 50%				
EB Traffic	16.09	0.99	16.17	0.99	18.4 *	0.71 *	12.65 *	0.65 *
NB Traffic	33.92	0.98	33.91	0.98	5.8 *	0.16 *	6.23 *	0.19 *
Crossing EB	25.32	0.82	25.34	0.82	12.17 *	0.56 *	10.58 *	0.54 *
Crossing NB	12.77	0.54	13.05	0.55	11.8 *	0.57 *	10.45 *	0.54
				$EB_{v/c}$ 95% - $NB_{v/c}$ 75%				
EB Traffic	19.62	1.23	19.81	1.25	18.5	0.72 *	10.74 *	0.54 *
NB Traffic	25.15	1.00	25.59	1.04	10.00 *	0.33 *	8.33 *	0.28 *
Crossing EB	22.6	0.78	22.62	0.78	11.94 *	0.56 *	10.18 *	0.52 *
Crossing NB	17.69	0.68	17.84	0.69	11.62 *	0.56 *	10.92 *	0.54 *
				$EB_{v/c}$ 95% - $NB_{v/c}$ 95%				
EB Traffic	26.97	1.75	27.82	1.82	17.30 *	0.63 *	9.72 *	0.44 *
NB Traffic	27.14	1.87	30.16	2.18	15.00 *	0.53 *	15.21 *	0.57 *
Crossing EB	22.71	0.79	22.72	0.79	11.95 *	0.54 *	10.95 *	0.54 *
Crossing NB	22.43	0.78	22.49	0.78	11.76 *	0.55 *	10.32 *	0.52 *

* Statistically significant at 99% confidence level compared to case (1), pre-timed with no active TSP

^ Statistically significant at 95% confidence level compared to case (1), pre-timed with no active TSP

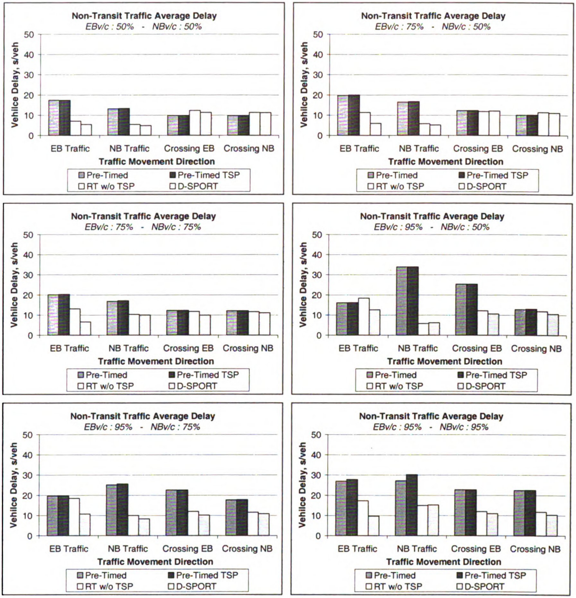


Figure 9.12 Simulated non-transit traffic delay for different v/c combinations and a boarding rate of 150 passenger/hr

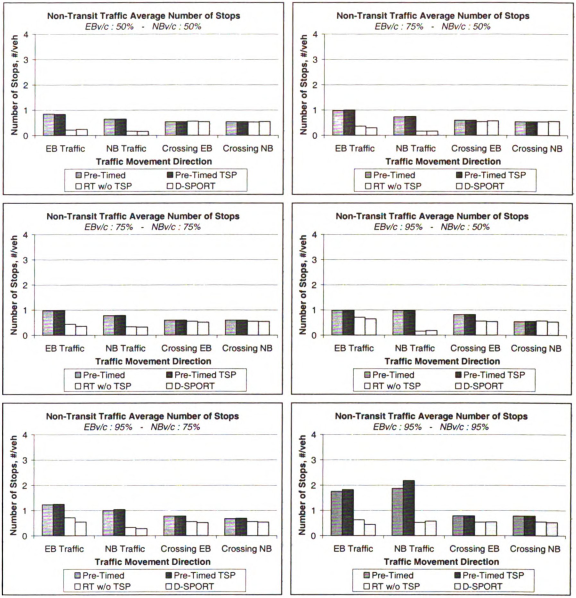


Figure 9.13 Simulated non-transit number of stops for different v/c combinations and a boarding rate of 150 passenger/hr

9.4. Schedule Adherence and Transit Service Reliability

Improving the schedule adherence and service reliability is part of the D-SPORT controller cost function. In order to quantify the service reliability for transit service users, transit lateness relative to the scheduled arrival for each transit vehicle is recorded in each simulation experiment. As defined in Section 6.3.3, the lateness is computed for each transit vehicle as the difference between the actual and scheduled arrival times at each bus stop, averaged across all the simulated buses in the 2-hr simulation replications. For a given bus stop, the scheduled arrival is calculated as the transit headway plus the expected travel time from the moment a transit vehicle enters the network until it reaches the bus stop while traveling at a desired speed of 35 mph (55 km/hr) without any impedance. For each scenario, the lateness is then averaged across all the transit vehicles that have dwelled at a bus stop over the 2-hr simulation duration. The standard deviation for lateness at given bus stop is finally computed based on all the recorded transit lateness over all the replications associated to a given scenario.

Table 9.5 summarizes the average and standard deviation of transit lateness for each simulated bus stop in each of the scenarios considered. The results show that the D-SPORT controller generally produced the lowest transit lateness compared to the other controllers, with average lateness varying between 30 to 80 seconds. Compared to the other controllers, lateness was reduced by 5% to 55% in most of the cases. Furthermore, variations in lateness from implementing the D-SPORT controller are lower by 50%-90%, with lateness standard deviations varying between 4.2 and 8.8 seconds. In this case, a low value of standard deviation means a better and more reliable service for passengers because of the lower probability of observing late transit arrivals at bus stops.

The results further show no statistical difference at a 95% confidence level between the pre-timed scenarios with or without active TSP. In this case, there is no statistical evidence to suggest that implementing TSP plans within pre-timed controllers would yield any significant improvement on schedule adherence. This finding is associated with the fact that active TSP is implemented locally while the MOEs are measured at the corridor and the network level. Active TSP does not take into account the projected arrival of a transit vehicle at the downstream intersection. Because of this, priority treatments granted at one intersection may result in more delay before it reaches the downstream bus stop because such an action could disturb traffic progression pattern.

Furthermore, the simulation results show that application of the D-SPORT controller reduced the average lateness of transit vehicles at bus stops when compared to real-time controller without TSP. This is due to the fact that the disabling the TSP sub-cost function removes the ability of the controller to keep track of the trajectory of transit vehicles. Although the real-time controller without TSP actions controller has resulted in comparable transit delays from the D-SPORT controller in some experiments, the D-SPORT controller provides a better transit service reliability for these scenarios. In terms of service reliability, the results of Table 9.5 also show a relatively lower lateness standard deviation resulted from implementing the D-SPORT controller, which means a more reliable schedule to the transit riders and the service users.

Table 9.5 **Transit lateness at bus stops for different traffic demand and a boarding rate of 50 passenger/hr**

Average Transit Lateness, Seconds									
Scenario		**Pre- Timed (1)**		**Pre- Timed TSP (2)**		**Rea-Time w/o TSP (3)**		**D-SPORT (4)**	
$EB_{V/C} - NB_{V/C}$		Lateness *s*	Std Dev. *s*	Lateness *s*	Std Dev. *s*	Lateness *s*	Std Dev. *s*	Lateness *s*	Std Dev. *s*
Bus stop 1	**50 - 50**	58.09	3.96	57.41	2.35	58.09	3.96	28.25 *	5.82
	75 - 50	66.80	11.92	66.68	11.81	66.8	11.92	32.67 *	5.88
	75 - 75	66.73	12.06	66.85	12.25	66.73	12.06	30.96 *	5.57
	95 - 50	70.94	13.92	71.12	14.25	70.94	13.92	57.79 *	5.55
	95 - 75	88.83	9.42	88.88	9.47	88.83	9.42	57.04 *	6.04
	95 - 95	86.15	6.39	86.34	5.52	86.14	6.41	54.00 *	4.20
Bus stop 2	**50 - 50**	77.86	3.58	77.73	1.58	77.86	3.58	51.71 *	6.55
	75 - 50	90.93	16.65	91.30	17.03	90.93	16.65	55.38 *	6.15
	75 - 75	90.93	16.71	91.58	17.06	90.93	16.71	51.13 *	7.32
	95 - 50	84.80	13.74	90.24 *	16.02	84.80	13.74	80.17 *	7.05
	95 - 75	97.63	12.89	100.25	16.18	97.63	12.89	79.88 *	8.81
	95 - 95	98.44	16.15	113.00 *	21.98	98.65	16.50	76.79 *	5.57
Bus stop 3	**50 - 50**	93.48	23.98	94.75	23.13	93.48	23.98	55.67 *	7.88
	75 - 50	91.57	16.56	88.71	16.67	91.57	16.56	54.17 *	7.62
	75 - 75	91.32	17.61	92.68	17.67	91.32	17.61	59.46 *	7.38
	95 - 50	112.70	20.25	88.27 *	22.30	112.70	20.25	57.58 *	6.46
	95 - 75	98.47	20.06	87.99 *	21.54	98.47	20.06	60.88 *	8.19
	95 - 95	101.86	16.70	109.16 ^	27.39	101.08	16.35	80.25 *	6.76
Bus stop 4	**50 - 50**	71.58	16.62	72.68	16.07	71.58	16.62	31.96 *	5.97
	75 - 50	66.60	12.51	64.05	12.09	66.60	12.51	31.46 *	5.33
	75 - 75	67.22	12.20	67.97	12.29	67.22	12.20	35.25 *	5.66
	95 - 50	72.45	16.58	64.90 *	12.01	72.45	16.58	31.83 *	6.59
	95 - 75	70.24	3.76	67.67 *	3.94	70.24	3.76	36.50 *	5.35
	95 - 95	74.12	16.94	75.03	17.79	73.58	16.08	59.71 *	5.48

* Statistically significant at 99% confidence level compared to case (1), pre-timed with no active TSP

^ Statistically significant at 95% confidence level compared to case (1), pre-timed with no active TSP

Figure 9.14 through Figure 9.17 summarizes the transit lateness fluctuation over time at each bus stop for different traffic demand levels and a transit demand of 50 passenger/hr. The results show that D-SPORT has performed better than the other controllers in most of the experiments. Most of the transit vehicles observed during the 2-hr simulation period have lower lateness compared to other experiments, with lateness standard deviation not more than 8.8 seconds. For the other controllers, the figures show a cyclic pattern in transit lateness over time. This cyclic pattern is due to the fact that transit lateness values are auto-correlated and not totally independent from each other. In this case, an early transit vehicle arrival at a bus stop may influence the arrival of other transit vehicles. For instance, fewer passengers might board the first vehicle as a result of its early arrival. This may then leave more passengers to board the next arriving vehicle, causing it to spend more time dwelling and to potentially being late arriving at its next stop. As can be seen in the figures, the D-SPORT controller was able to reduce the transit lateness auto-correlation by reducing the transit lateness variance, regulating the transit service, and improving the transit schedule adherence.

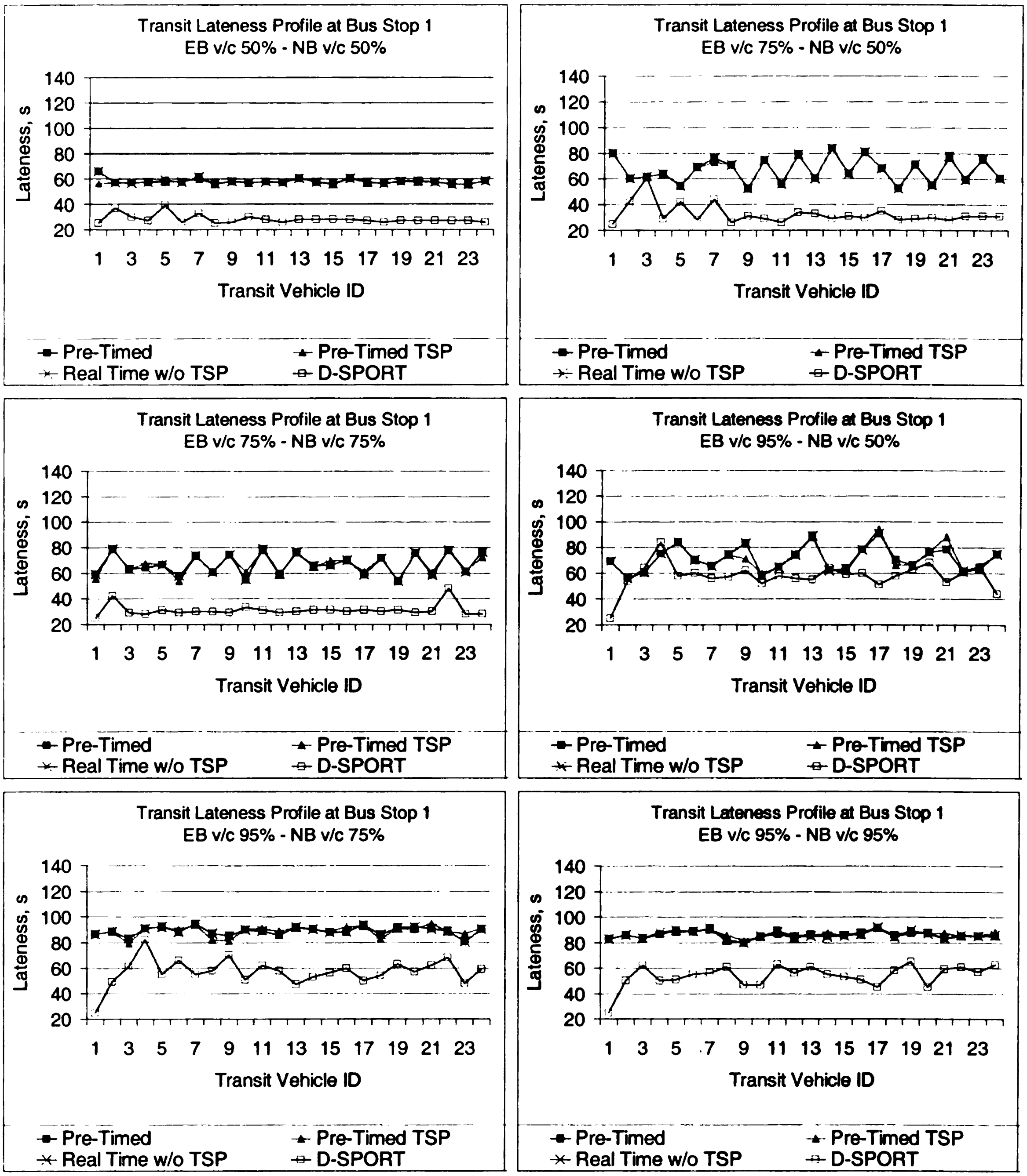


Figure 9.14 Transit lateness at bus stop 1 for scenarios and 50 passenger/hr boarding rate

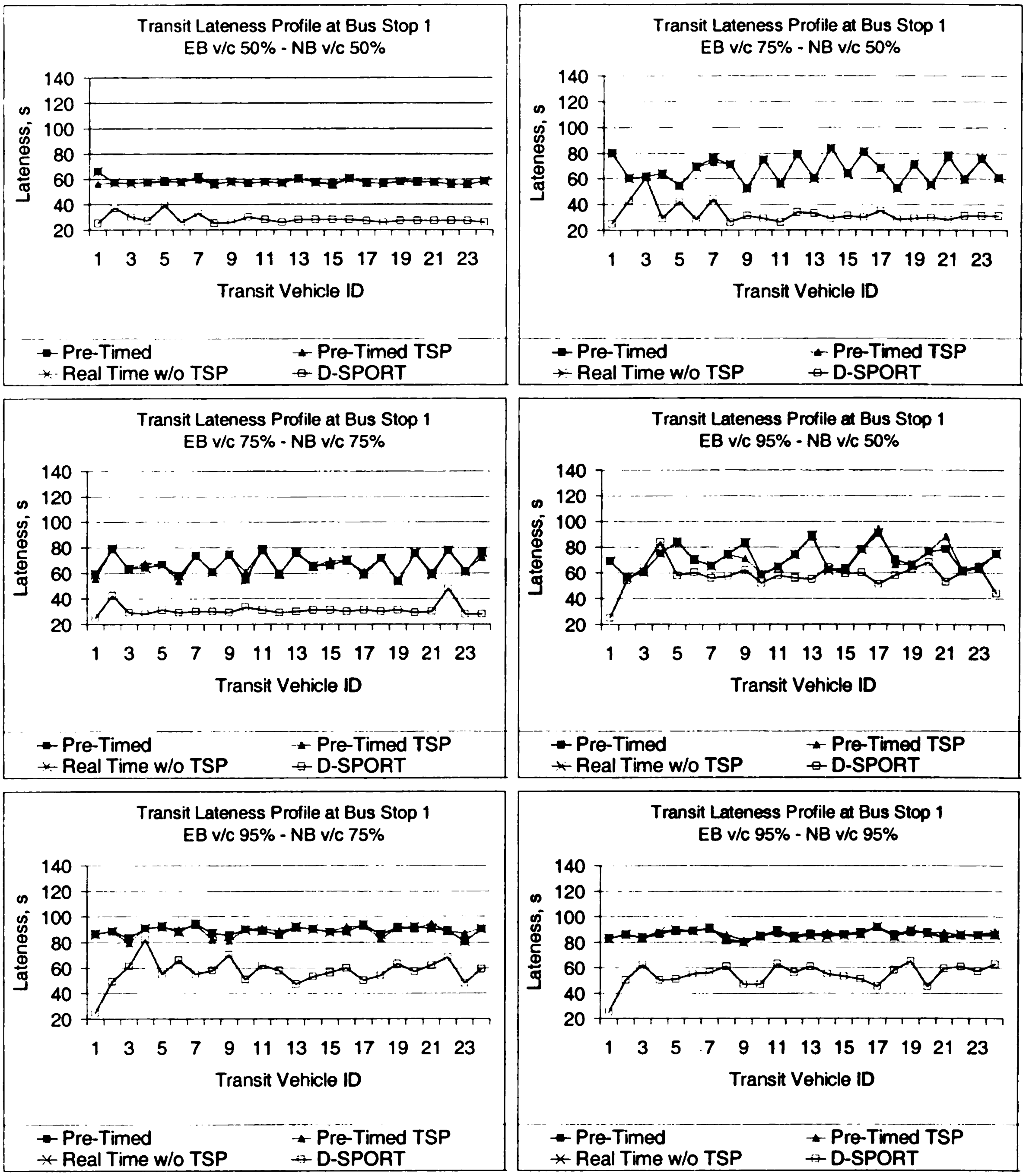


Figure 9.14 Transit lateness at bus stop 1 for scenarios and 50 passenger/hr boarding rate

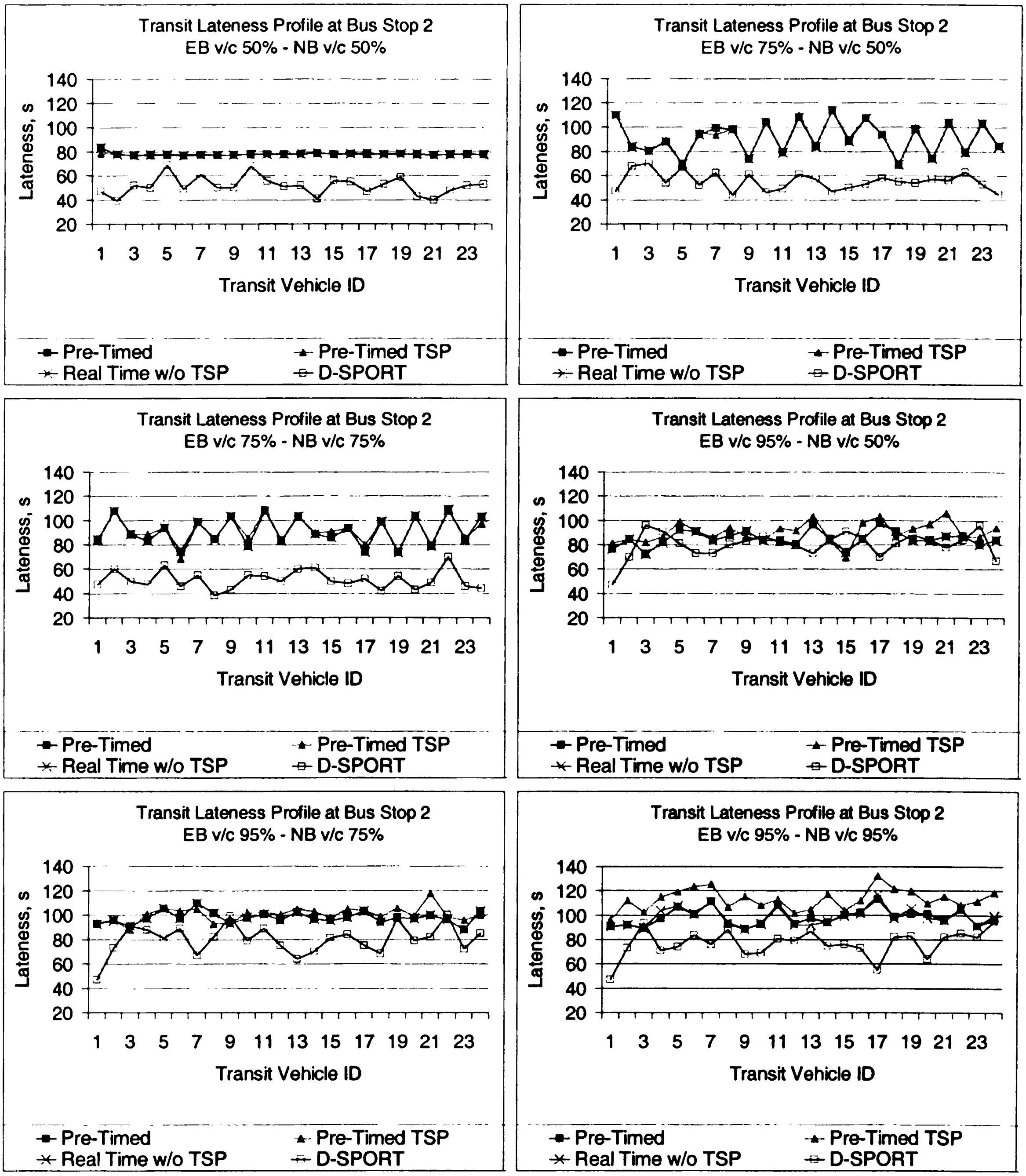


Figure 9.15 Transit lateness at bus stop 2 for scenarios and 50 passenger/hr boarding rate

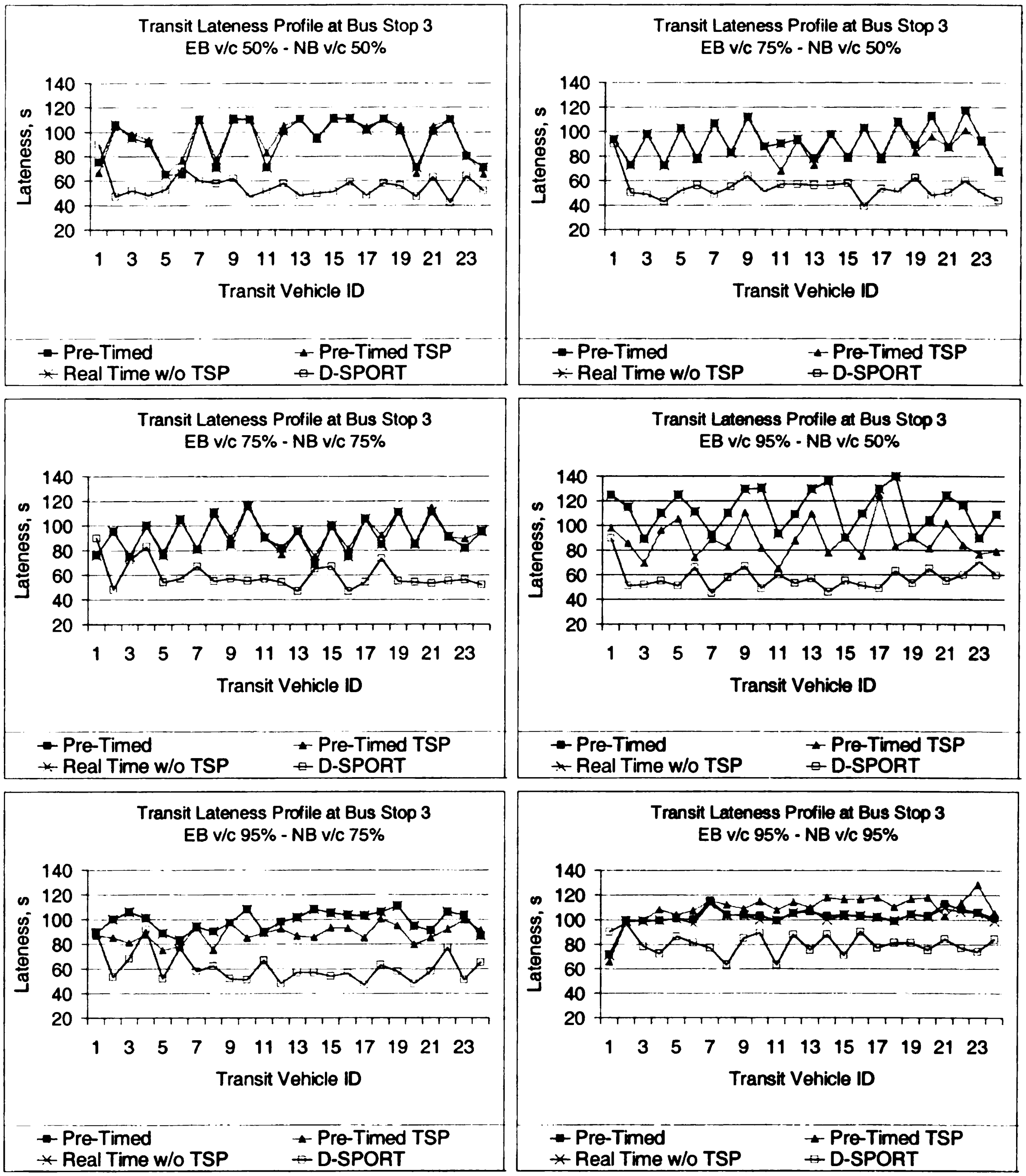


Figure 9.16 Transit lateness at bus stop 3 for scenarios and 50 passenger/hr boarding rate

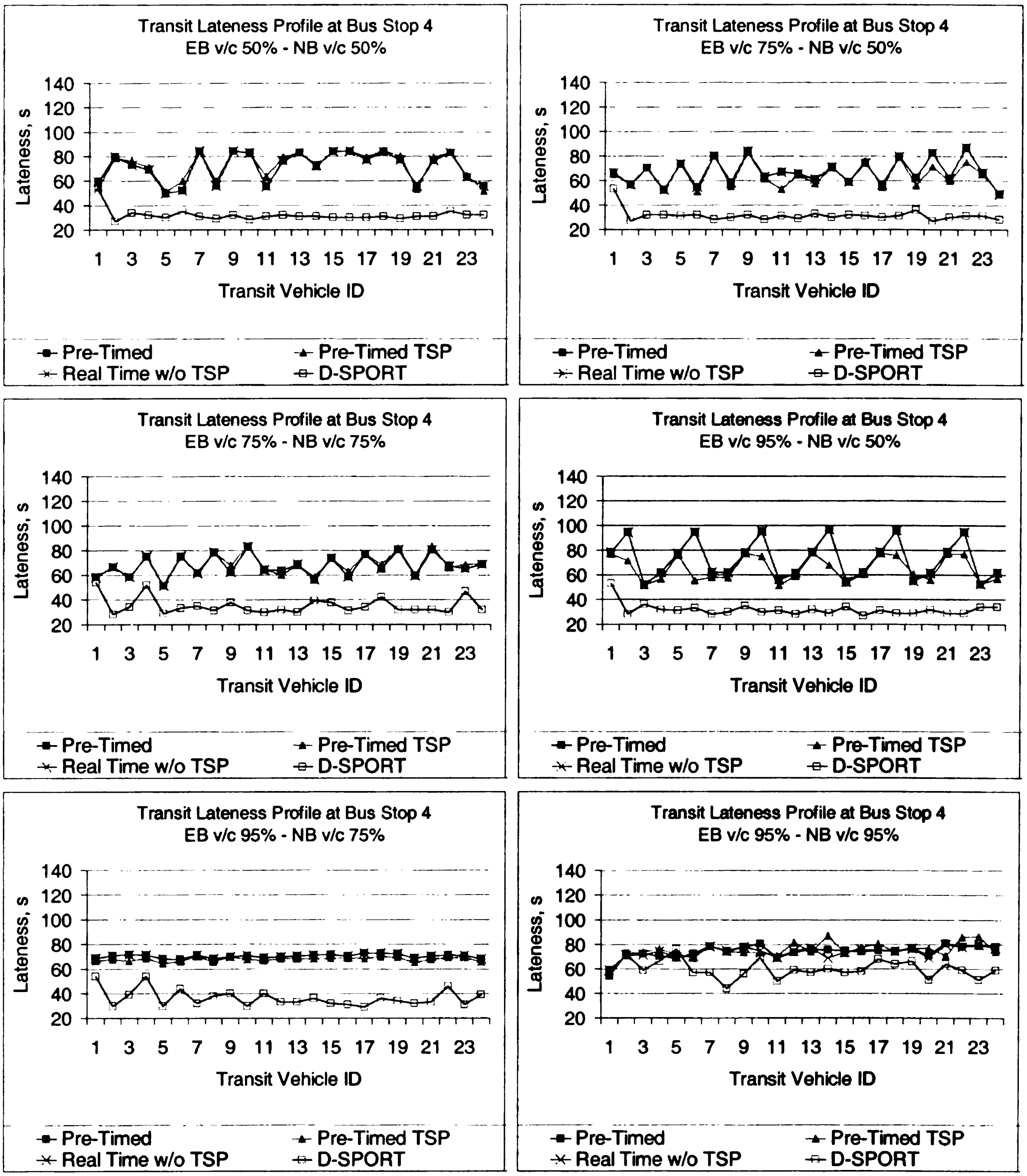


Figure 9.17 Transit lateness at bus stop 4 for scenarios and 50 passenger/hr boarding rate

9.5. Summary

This chapter summarizes the results of simulation experiments that were performed to assess traffic network performances under different traffic control approaches in a range of traffic and transit demand scenarios. The results show a significant reduction in transit delay and number of stops when the D-SPORT controller is implemented. The results further show that D-SPORT is able to react to traffic fluctuations by allocating green times at traffic signals along the network to accommodate both general traffic and transit movements. The simulation results further show that the D-SPORT controller has the ability to improve and regulate transit schedule adherence at bus stops by reducing transit lateness, as well as to improve transit service reliability by reducing the variation of transit arrivals at bus stops.

CHAPTER X.

CONCLUSIONS AND RECOMMENDATIONS

Efforts were made in this research to develop and integrate TSP and a real-time traffic signal control model for an arterial network. The importance of this work is to provide practical solutions for several issues related to the integration of the needs of both transit vehicles and general traffic. This chapter summarizes the work performed in this research, the major findings of the research, and recommendations for future work.

10.1. Introduction

Implementing TSP in coordinated networks of signalized arterials presents several issues linked to the differing operational requirements associated with transit vehicles and the general traffic. The primary problem is that transit vehicles often do not maintain progression with the general traffic stream due to their need for accommodating passenger boarding and alighting at bus stops. Another issue is related to the nature of TSP operations. TSP operations aim to provide preferential treatments to transit vehicles approaching individual intersections. While granted preferential treatments at an intersection may benefit transit vehicles, such actions may negatively impact local traffic and have further negative repercussions at other surrounding intersections if network-based operational objectives are not considered. For instance, providing transit vehicles with preferential treatment might lead to an increased reliability in schedule adherence, which would be an obvious benefit to transit riders. However, temporary changes in signal timings could also interfere with the established traffic progression patterns. An ensuing loss of progression could potentially increase delay at downstream intersections for non-transit vehicles, thus potentially negating all the transit benefits from a system standpoint. Furthermore, the reception of conflicting or simultaneous priority requests is generally not handled appropriately in existing systems. Priority requests are handled on a first-come-first-serve basis. Better results could be achieved if attempts were made to better utilize potentially available information about transit operations and traffic conditions.

The objective of this research was to develop a dynamic signal priority optimization in real-time traffic control system (D-SPORT), which is capable of providing efficient transit priority at signalized intersections operated within a coordinated network. Efficiency was defined as the ability to benefit transit vehicles, in terms of reducing transit delay and improving schedule adherence by reducing transit lateness, while improving non-transit traffic conditions by reducing traffic delay and number of stops.

To allow the proposed signal controller to consider all the performance measures of interest, a cost function was developed to combine all of the aforementioned objectives in one mathematical formula, and express them in relation to variables that can be measured by available traffic surveillance and vehicle tracking technologies, such as traffic volumes, transit vehicle location, and signal control timing parameters. Within this cost function, estimates of delays incurred by the general traffic are obtained by utilizing the delay equations defined in the HCM 2000, while the number of stops incurred by individual vehicles is computed using macroscopic queuing analysis models.

The D-SPORT controller bases its priority actions on the anticipated time a transit vehicle will need to reach a specific point along its route (such as a bus stop or a stop line). This travel time is a function of the distance between a transit vehicle's current location and the selected target point, the desired transit speed, the expected transit delay before reaching the target point, and the time consumed to board and alight passengers at bus stops located between their current location and the target point are both considered as well. Similar to traffic performance parameters, the expected transit delay is calculated based on HCM 2000 delay equations. The prediction of dwell time at bus stops is on the other hand achieved using an ANN model that was developed in this research and which was demonstrated to produce good results.

The determination of appropriate signal timing actions is finally achieved using a GA model seeking to minimize the cost function value. The value of the cost function is a function of delay, the number of stops, transit travel time, and transit vehicle schedule.

Testing of the D-SPORT controller was performed using a virtual network coded in the VISSIM microscopic simulation model. The tests examined the performance of the proposed signal control logic compared to alternative traffic signal control approaches commonly used in practice. Performance of the proposed model was compared more specifically with (1) pre-timed signal control optimized offline using SYNCHRO 6, (2) pre-timed signal control with active TSP offering green extensions and early green recalls, and (3) a real-time signal control that does not provide any preferential treatments for transit vehicles.

10.2. Major Conclusions and Research Contributions

The major findings of this research are summarized below:

- Dwell time variability affects signal controller's ability to provide transit vehicles with efficient priority treatments. The danger is for the controller to provide a specific treatment that will end up not being used because of a longer or shorter than expected dwell time. Such wasted treatments could negatively affect traffic signal control at the prioritized intersection and cause unduly increases in vehicle delays. This is particularly significant issue at intersections with near-side bus stops, where buses stopping near the stop line provide a limited timeframe to implement TSP strategies following indication that a bus is ready to leave.

- Although dwell time variability might impact the ability of transit vehicles to benefit from priority systems, dwell time variability can be accommodated by changing transit vehicle arrival patterns at an intersection relative to the traffic signal state. The primary action would be to increase the percentage of transit vehicles arriving within a time window for which they is a higher probability to benefit from an early green recall or a green time extension, i.e., less sensitivity to changes in actual arrival time.

- While the potential effects of dwell time variability in any TSP algorithm is an important key in transit priority planning and operations, its consideration depends on the objective of the study. If the objective is to better model and evaluate general traffic impacts associated with land development, or roadway improvements, then experiments using average dwell time can be used. If the objective is instead to determine the efficiency of selected priority strategies or algorithms with respect to transit operations, or the efficiency of signal control with TSP treatment, it then becomes important to consider dwell time variability in the TSP optimization.

- An ANN prediction model was successfully used to predict the stop line arrival time of buses traveling on an intersection with near-side bus stops. Results of model testing show an ability to improve the operation of transit signal priority systems by enhancing the ability to correctly predict bus arrival times at intersections and reduce unnecessary signal timing alterations. In the test scenarios considered, this enhanced prediction capability resulted in lower delay for buses and lower impacts on general traffic when compared to a priority system simply basing its decisions on average dwell and travel times.

- The GA was successfully able to optimize multiple cost functions concurrently and to return a near-optimal signal timings for a real-time traffic control system to accommodate both transit and non-transit traffic streams at the traffic network level. The use of the GA has strongly promoted the real-time implementation of the D-SPORT controller in the field due to the short time that is required to perform the optimization process.

- The performance of the D-SPORT controller was examined under different traffic and transit demands levels, and was compared to other common-used signal control systems. The simulation experiments have resulted in a significant reduction in transit delays and number of stops. The simulation results also show the D-SPORT controller ability to react to traffic fluctuations and to re-distribute the green times in the network to accommodate general transit and non-transit needs. In terms of schedule adherence, the D-SPORT implementation reduced both transit lateness at bus stops and variations in bus stop arrival times.

- The D-SPORT controller was able to solve several issues related to implementing TSP in a coordinated traffic system. More specifically, the D-SPORT controller is insensitive to the bus stop classifications, the variability of dwell times and its impact on accurately predicting transit travel times, implementing TSP in a coordinated traffic networks, and the issue of simultaneous priority recalls.

10.3. Recommendations for Future Research

Although this research has addressed, discussed, and solved several issues related to integrating TSP within a real-time signal control coordinated traffic network, other issues can be addressed and examined. The following summarizes some of the most important issues that might need further and more detailed investigation:

- Further investigations to improve the proposed model applicability by taking into account more complicated network layouts and geometric designs such as grid networks, phasing schemes, turning volume, traffic composition, saturated traffic conditions, and two-way traffic networks.

- Considering the pedestrian movements for more comprehensive modeling that balances the need of transit vehicles, the general traffic, and pedestrians.

- Sensitivity analysis of the D-SPORT controller cost functions to the microscopic delay, queue and number of stops models.

- The consideration of a car-following model in developing the cost function, and determining the D-SPORT controller efficiency and the traffic network performance.

- Adapt the dynamic speed concept as a part of the D-SPORT cost function.

- Determine the D-SPORT controller's ability to respond to unexpected traffic events, such as traffic incidents, and the effect of sink-and-source traffic movements.

- Actual field experiments to validate the applicability of the D-SPORT controller to real-world conditions, where other factors, such as weather and difficult to model driver behavior, might influence the traffic network performance.

REFERENCES

1. Baker et al., *An Overview of Transit Signal Priority*. 2002, Washington, D.C.: ITS America.

2. Vincent, R.A.a.H., K., *Public Transport Priority as a Signalized-Controled Junction: An Experiment in Bern, Switzerland.* Traffic Engineering & Control, 1970.

3. Teng, H., et al., *Simulation Testing of Adaptive Control, Bus Priority, and Emergency Vehicle Preemption in New York City*. in *Transportation Reseach Board, 82nd Annual Meeting*. 2003. Washington, D.C.

4. Wattleworth, J.A., K.G. Courage, and C.E. Wallace, *Evaluation of Bus-Priority Strategies on Northwest Seventh Avenue in Miami.* Transportation Research Record, 1977(626): p. 32-35.

5. Klein, L.A., *Sensor Technologies and Data required for ITS, book.* 2001, London: Artech House Publishers.

6. Parviainen, A., *An Overview of available and Developing Highway Vehicle Electronic Technology and Research Program*. Transportation Technology and Energy Branch, Ministry of Transportation. 1990, Ontario, Canada.

7. Drane, C. and C. Rizos, *Positioning Systems in Intelligent Transportation systems*. 1997, London: Artech House Publishers.

8. Pearson, R. *Traffic Signal Control.* [cited 3/15/2007]; Available from: http://www.calccit.org/itsdecision/serv_and_tech/Traffic_signal_control/traffsigrep_print.htm#adaptive.

9. Bretherton, R.D., N.B. Hounsell, and B. Radia. *Public Transport Priority in SCOOT.* in *Proceedings of the 3rd World Congress on Intelligent Transport systems.* 1996.

10. Cranshaw, V. and M.J. Shaw. *Vehicle Detection in Norwich* in *Highways and Transportation.* 1995.

11. Mathan H. Gartner, *OPAC: A Demand Responsive Strategy for Traffic Signal Control.* Transportation Research Record, TRR, National Research Council, Washington, D.C., 1983. 906.

12. Jacobson, K.L. *Transit Signal Priority Treatments in the Puget Sound Region.* in *Pacific Rim Transtech Conference.* 1993. Seattle.

13. Donati, f., et al., *A Hierarchical decentralized traffic light control system, the first realization.* IFAC 9th World congress, 1984. 11(11G/A-1).

14. Head, K.L. and P.B. Mirchandani, *A real-time traffic signal control system: architecture, algorithms, and analysis.* Transportation Research. Part C: Emerging Technologies, 2001. 9(6): p. 415-432.

15. Head, K.L., P.B. Mirchandani, and S. Shelby, *Estimation and prediction in a real-time traffic adaptive signal control logic: the RHODES prototype.* Informs meeting, San Diego, 1997.

16. Mirchandani, P.B. and D.E. Lucas, *Integrated Transit Priority and Rail/Emergency Preemption in Real-Time Traffic Adaptive Signal Control.* Intelligent Transportation systems, 2004. 8: p. 101-115.

17. Dion, F., H. Rakha, and Y. Zhang., *Evaluation of Potential Transit Signal Priority Benefits along a Fixed-Time Signalized Arterial.* Journal of Transportation Engineering, 2004. 130(3): p. 294-303.

18. Deshpande, V. and J. Collura. *ransit Signal Priority: Green Extension Benefits in Congested Corridor.* in *Preprint CD-ROM, 83rd Annual Meeting of the Transportation Research Board.* 2003. Washington, D.C.

19. Ngan, V., T. Sayed, and A. Abdelfatah, *Impacts of Various Traffic Parameters on Transit Signal Priority Effectiveness.* Journal of Public Transportation, 2004. 7(3): p. 71-93.

20. Chang, G.-L., M. Vasudevan, and C.C. Su, *Bus Preemption under Adaptive Control Environments.* Transportation Research Record 1494, TRB, National Research Council, Washington, D.C., 1995: p. 146-154.

21. *Transit Capacity and Quality of Service Manual, 2nd Edition.* . TCRP Report 100, ed. T.C.R. Program. 2003, Washington, D.C.: TRB, National Research Council.

22. TOWNES, M.S., et al. *TCRP Report 19. Guidelines for the Location and Design of Bus Stops.* 1996, Washington, D.C.: Transportation Research Board, Transportation Research Board.

23. Huffman, B., et al. *Signal Priority Demonstration in Suburban Chicago.* in *Proceedings, 1998 Bus Operations, Technology and Management Conference.* May 1998.

24. Chada, S. and R. Newland, *Effectiveness of Bus Signal Priority. Report NCTR-416-04.* National Center for Transit Research, University of South Florida. 2002, Tampa, FL.

25. Feng, Y., J. Perrin, and P.T. Martin. *Bus Priority of SCOOT Evaluated in a VISSIM Simulation Environment.* in *Preprint CD-ROM, 82nd Annual Meeting of the Transportation Research Board.* 2003. Washington, D.C.

26. Byrne, N., et al. *Using Hardware-in-the-Loop Simulation to Evaluate Signal Control Strategies for Transit Signal Priority.* . in *Preprint CD-ROM, 84th Annual Meeting of the Transportation Research Board.* 2005. Washington, D.C.

27. Ghanim, M., F. Dion, and G. Abu-Lebdeh. *Artificial Neural Network Bus Arrival Time Prediction Tool for Transit Signal Priority with Nearside Bus Stop.* in *86th Annual Meeting of the Transportation Research Board.* 2007. Washington, D.C.

28. Dion, F. and M. Ghanim. *Impact of Dwell Time Variability on Transit Signal Priority Performance at Intersections with Nearside Bus Stop.* in *86th Annual Meeting of the Transportation Research Board.* 2007. Washington, D.C.

29. Kim, W. and L. Rillett. *An Improved Transit Signal Priority System for Networks with Nearside Bus Stops.* in *Preprint CD-ROM, 84rd Annual Meeting of the Transportation Research Board.* 2005. Washington, D.C.

30. Rakham, H. and Y. Zhang, *Sensitivity Analysis of Transit Signal Priority Impacts on the Operation of a Signalized Intersection.* Journal of Transportation Engineering, 2004. 130(6): p. 796-804.

31. Chien, S., Y. Ding, and C. Wei, *Dynamic Bus Arrival Time Prediction with Artificial Neural Networks.* Journal of Transportation Engineering, 2002. 128(5): p. 429-437.

32. Al-Sahili, K.A., and Taylor, W.C. , *Evaluation of Bus Priority Signal Strategies in Ann Arbor, Michigan.* Transportation Research Record 1554, Transportation Research Board, Washington, D.C., 1995: p. 74-79.

33. Duerr, P.A., *Dynamic Right-of-Way for Transit Vehicles: An Integrated Modeling Approach for Optimizing Signal Control on Mixed Traffic Arterials.* Transportation Research Record, TRB, National Research Council, Washington, D.C., 2000. 1731: p. 31-39.

34. Sun, W. and Y.-x. Yuan, *Optimization Theory and Methods, Nonlinear Programming*. Vol. 1. 2006, New York: Springs.

35. Robinett, R.D., et al., *Applied Dynamic Programming for Optimization of Dynamic Systems`*. 2005, Philadelphia, PA: Society for Industrial and Applied Mathematics.

36. Murty, K.G., *Operations Research: Deterministic Optimization Models*. 1995, Englewood Cliffs, NJ: Prentice Hall.

37. Holland, J.H., *Adaptation in Artificial and Natural Systems*. University of Michigan Press. 1975, Ann Arbor.

38. Goldberg, D.E., *Genetic Algorithms in Search, Optimization, and Machine Learning*. 1989, Boston, MA, USA Addison-Wesley Longman Publishing Co., Inc.

39. Ayyub, B.M. and R.H. McCuen, *Probability, Statistics, & Reliability for Engineers*. 1997, New York: CRS Press.

40. Haykin, S., *Neural Networks: A Comprehensive Foundation*. 1994, New York, NY: Macmillan.

41. Singh, A. and G. Abu-Lebdeh, *Conditional Independence Graph-based State-Space Neural Network Modeling for Short-Term Travel Time Prediction on Urban Arterials.* Presented at Artificial Neural Networks in Engineering (ANNIE), St. Louis, MO, 2006.

42. Mandavilli, S., Y Najjar, and G. Abu-Lebdeh, *Modeling Crash Rates For the Kansas Rural Expressway Network.* Artificial Neural Networks in Engineering (ANNIE), 2005. 15: p. 721-730.

43. Ahmed, K., G. Abu-Lebdeh, and R.W. Lyles, *Prediction of Pavement Distress Index with Limited Data on Causal Factors: An Autoregression Approach.* International Journal of Pavement Engineering, 2006. 7(1): p. 23-35.

44. Bailer-Jones, D. and M. Bailer-Jones. *C.A.L. Modelling data: Analogies in neural networks, simulated annealing and genetic algorithms. In Model-based reasoning: Scientific discovery, technological innovation, values, Kluwer/Plenum., 2001.* in *Model-based reasoning: Scientific discovery, technological innovation, values*. 2001. Kluwer/Plenum.

45. May, A.D., *Traffic Flow Theory*. 1990, New Jersy: Prentice-Hall.

46. *CORSIM user manual version 1.01.* August 1996: Federal Highway Administration.

47. *AIMSUN 2 User Manual.* Vol. 4.0. Novermber 2000: Transport Simulation systems.

48. *Paramics traffic simulation modeller V3.0: Reference manual.* 1999, Edinburgh: Quadstone Ltd.

49. Planung, P., *VISSIM 4.10 Users Manual.* PTV Planung Transport Verkehr AG. 2005, Karlsruhe, Germany.

50. Chapman, S.J., *Fortran 90/95 for Scientists and Engineers.* 1 ed. 1998, Boston, USA: McGraw-Hill Companies.

51. KIM, W., *An Improved Bus Signal Priority System for Networks with Nearside Bus Stops*, in *Office of Graduate Studies.* 2004, Texas A&M University.

52. Trafficware. *SYNCHRO 6. Traffic Signal Simulation Software User Manual.* 2005. Albany, CA.

53. Rakha, H., M. Snare, and F. Dion, *Vehicle Dynamics Model for Estimating Maximum Light Duty Vehicle Acceleration Levels.* Transportation Research Record, 2004. 1883: p. 40-49.

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

55. *Highway Capacity Manual, Sepcial Report 209*, in *Transportation Research Board.* 2000, National Research Council: Washington, D.C.

56. Garber, N.J. and L.A. Hoel, *Traffic and Highway Engineering.* Third ed. 2002, Pacific Grove, CA: Broose/Cole.

57. Kang, Y.-S., *Delay, Stop and Queue Estimation for Uniform and Random Traffic Arrivals at Fixed-Time Signalized Intersections*, in *Civil and Environmental Engineering.* 2000, Virginia Polytechnic Institute and State University.

58. Planung, P., *VISSIM 4.30-00 COM Interface Manual.* PTV Planung Transport Verkehr AG. 2007, Karlsruhe, Germany.

59. Schizas, C.N., et al., *Artificial Neural Networks in Forecasting Minimum Temperature.* Second International Conference on Artificial Neural Networks, 1991(18-20 Nov 1991): p. 112-114.

60. Battiti, R., *First-and second-order methods for learning: between steepest descent and Newton's method.* Neural Computation, 1992. 4(2): p. 141-166.

61. Hagen, M.T. and M.B. Menhaj, *Training feedforward networks with the Levenberg-Marquardt algorithm.* IEEE Trans. on Neural Networks, 1994. 5(6): p. 989-993.

62. Demuth, H. and M. Beale, *Neural Network Toolbox For Use with MATLAB, User's Guide*. Vol. 3. 1998, Natick, MA: The MathWorks.

63. Moreira, J.E., S.P. Midkiff, and M. Gupta, *A comparison of Java, C/C++, and FORTRAN for numerical computing.* Antennas and Propagation Magazine, IEEE, 1998. 40(5): p. 102-105.

APPENDIX A: ANN MODEL SOURCE CODE IN MATLAB 6

```
%with normalization tansig purelin,50-3
% Size of original data m is the number of cases and n is the number of
% attributes ( including the three targets)
[m,n]=size(all);
disp('Original Data dimensions [Attributes Cases]');
disp([n m]);

m_dim = m;
n_dim = n;


% Normalization
p = all(:,4:17); p = p';
t = all(:,1); t = t';
[pn,minp,maxp,tn,mint,maxt] = premnmx(p,t);

% Preparing data for Training, Validation and Test, 1/2 data for
training
% and 1/4 for validation and 1/4 to test

iitr =  1:1:5000;
iival = 5001:1:7000;
iitst = 7001:1:m;

val.P = pn(:,iival); val.T = tn(:,iival);
test.P = pn(:,iitst); test.T = tn(:,iitst);
ptr = pn(:,iitr); ttr = tn(:,iitr);
```

```
% Setting up Lambert-Marrquardt FF Backpropagation Network. (More
details at 'Neural Network Toolbox: Backpropagation: Levenberg-
Marquardt (trainlm)')
disp('Setting up network ...');

% Network Architecture and Sigmoid Functions
net = newff(minmax(ptr), [8  1],{'logsig' 'purelin'},'trainlm');



% Training Parameters
net.trainParam.epochs = 500;
net.trainParam.goal = 1e-4;          % MSE to reach 0.00000001
net.trainParam.max_fail = 10 ;        % Five Validations to check when
error rises
net.trainParam.mem_reduc = 1;        % Full Jacobian Calculated, no
memory restrictions
net.trainParam.min_grad = 1e-8;
net.trainParam.mu = 0.0001;
net.trainParam.mu_dec = 0.1;
net.trainParam.mu_inc = 10;
net.trainParam.mu_max = 1e10;
net.trainParam.show = 1;
net.trainParam.time = inf;

% Learning Rate
net.trainparam.lr = 0.1;


% Performance = 'Mean Square Error'
net.performFcn = 'mse';



% Training
disp('Training ...');
[net,tr]=train(net,ptr,ttr,[],[],val,test);

% Ploting the training results
figure(10);
title('Plot of Training, Validation and Test by Epochs'); hold on
plot(tr.epoch,tr.perf,tr.epoch,tr.vperf,tr.epoch,tr.tperf)
legend('Training','Validation','Test',-1);
ylabel('Mean Squared Error'); xlabel('Epoch')
hold off

% Simulation and Post training analysis
an = sim(net,pn);
a = postmnmx(an,mint,maxt);


wts = net.iw;


% Regression Analysis
figure(1)
title('Regression Analysis of Results'); hold on
```

```
[m,b,r] = postreg(a(1,:),t(1,:)); hold off


% Results - Network Training
an2 = sim(net, ptr);
a2 = postmnmx(an2,mint,maxt);
figure(2);
title('Regression Analysis of Training'); hold on
[m,b,r] = postreg(a2(1,:),t(:,iitr)); hold off



% Results - Network Validation
an4 = sim(net, val.P);
a4 = postmnmx(an4,mint,maxt);
figure(4);
title('Regression Analysis of Validation'); hold on
[m,b,r] = postreg(a4(1,:),t(:,iival)); hold off


% Results - Network Test
an3 = sim(net, test.P);
a3 = postmnmx(an3,mint,maxt);
figure(3);
title('Regression Analysis of Testing'); hold on
[m,b,r] = postreg(a3(1,:),t(:,iitst)); hold off




format short;
disp('Network Architecture');
disp('Number of Layers');
disp(net.numlayers+1);
nneurons = 7;
slayer = size(net.layers);
for i=1:slayer(1,1)
    nneurons =[nneurons net.layers{i}.size];
end


predicted = a3';
actual = t(7000:m_dim)';
```

APPENDIX B: GA PLATFORM SOURCE CODE IN FORTRAN 90/95

```
!!!!!!!    Genetic Algorimth Code for Traffic Signal Controller USE.
!!!!!!!
!!!!!!!                        Written By:
!!!!!!!
!!!!!!!                 Mohammad Shareef Ghanim
!!!!!!!
!!!!!!!                 Civil Engineering Dept.
!!!!!!!                Michigan State University
!!!!!!!                     June 1st 2008
PROGRAM Genetic
IMPLICIT NONE

INTEGER, ALLOCATABLE, DIMENSION(:,:):: ipop, offspring, child
REAL, ALLOCATABLE, DIMENSION(:,:):: m_rnd
REAL:: s_rnd
INTEGER, ALLOCATABLE, DIMENSION(:):: parent
REAL, ALLOCATABLE, DIMENSION(:,:):: var, x, eval
REAL, ALLOCATABLE, DIMENSION(:,:):: bestever
REAL, ALLOCATABLE, DIMENSION(:):: randsel, minmax
INTEGER, ALLOCATABLE, DIMENSION(:):: counter
INTEGER:: cumcounter
INTEGER:: popsize, generation, dv, str
INTEGER:: bestindex, elite
INTEGER:: cross, mutate
INTEGER::i, n, j, igen
REAL:: str_mutate
INTEGER:: i_mutate
REAL:: minfit, maxfit, sumfit,avgfit
REAL:: minnorm, maxnorm, sumnorm, avgnorm
REAL:: best
REAL::cross_p, check_cross, mutate_p, check_mutate
REAL:: cross_i, mutate_i
```

```
REAL, ALLOCATABLE, DIMENSION(:):: fit, freq
REAL, ALLOCATABLE, DIMENSION(:)::nevals
INTEGER:: eval_count

REAL:: time_in, time_out
REAL:: s
REAL:: fitval
CHARACTER:: dummy
CHARACTER:: selection, check_elitism
INTEGER, DIMENSION(1:1):: rand_seed
INTEGER:: icase

eval_count = 1
! Constants
dv                  =     42
str                 =     6

ALLOCATE(minmax(1:2))
! Genetic Algorithm Parameters
OPEN ( UNIT = 100, FILE = 'C:\Final Project\Project
Network\ga_parameters.txt', STATUS = 'UNKNOWN')
READ(100, *) dummy, selection       ! Selection Method, R for Roulette
Wheel, T for Tournament Selection
READ(100, *) dummy, popsize         ! Population Size, Should be EVEN
number
READ(100, *) dummy, generation      ! Number of Generations
READ(100, *) dummy, cross_p         ! Crossover Probability
READ(100, *) dummy, mutate_p        ! Mutation Probability
READ(100, *) dummy, check_elitism   ! Applying Eltisim? 'Y' or 'N'
READ(100, *) dummy, rand_seed       ! Random seed
READ(100, *) dummy, minmax(1)       ! Minimum Green
READ(100, *) dummy, minmax(2)       ! Maximum Green
CLOSE(100)
!!!!!!!!!!!
! This is to activate MultiRun with different Random Seeds

icase = 100

call cpu_time(time_in)
rand_seed = icase
CALL RANDOM_SEED(put = rand_seed)

ALLOCATE(ipop(1:popsize,1:dv*str))
ALLOCATE(offspring(1:popsize,1:dv*str))
ALLOCATE(child(1:popsize,1:dv*str))
ALLOCATE(var(1:popsize, 1:dv))
ALLOCATE(x(1:popsize, 1:dv))
ALLOCATE(eval(1:popsize,1:3))
ALLOCATE(randsel(1:popsize))
ALLOCATE(parent(1:popsize))
ALLOCATE(bestever(1:generation, 1:2))
ALLOCATE(counter(1:generation))
ALLOCATE(fit(1:generation*popsize))
ALLOCATE(freq(1:generation*popsize))
ALLOCATE(nevals(1:generation))
counter = 0

!!!!!!!!!!!!!!!!!
!                 Boundary  Limits        !
!!!!!!!!!!!!!!!!!
!     minmax(1)=  20.
!     minmax(2)=  25.

!!!!!!!!!!!!!!!!!
```

```
!                  Generate Initial Strings      !
!!!!!!!!!!!!!!!!!
ALLOCATE(m_rnd(1:popsize, 1:dv*str))
CALL RANDOM_NUMBER(m_rnd)
      ipop = NINT(m_rnd)
DEALLOCATE(m_rnd)
!!!!!!!!!!!!!!!!!
!                  Starting The GA Process       !
!!!!!!!!!!!!!!!!!
DO igen = 1, generation
WRITE(*,*)'======================'
WRITE(*,*)'STARTING GENERATION NO: ', igen
WRITE(*,*)'======================'

WRITE(*,*)'...............................................'
!!!!!!!!!!!!!!!!!
!                  Decode the string to variables                    !
!!!!!!!!!!!!!!!!!
! WRITE(*,*)'================='
! WRITE(*,*)'Decoded Variables'
! WRITE(*,*)'================='
var=0.
x=0.
DO i = 1,popsize
  DO n = 1,dv
    DO j=1+(n-1)*str, str*n
      var(i,n)=var(i,n) + ipop(i,j) * ( 2.** (( str * n ) - j) )
    END DO
x(i,n) = (var(i,n)/(2.**str)) * ( minmax(2)- minmax(1) )+ minmax(1)
  END DO
END DO
!DO i = 1, popsize
!DO n = 1, dv
!x(i,n) = (var(i,n)/(2**str-1))*(minmax(2)-minmax(1))+minmax(1)
!END DO
!END DO

!!!!!!!!!!!!!!!!!
!                  Calculate the Fitness         !
!!!!!!!!!!!!!!!!!
! WRITE(*,*)'=================='
! WRITE(*,*)'Variables Fitness'
! WRITE(*,*)'=================='

! Rounding to the nearest 5-second green duration
x = NINT(x)
eval=0.
!Calculate Fitness of strings
DO i = 1,popsize
CALL OPT(x(i,:), fitval)
eval(i,1)=fitval
fit(eval_count) = fitval
fitval = 0.
eval_count = eval_count+1
END DO

sumfit=sum(eval(:,1))
avgfit=sumfit/popsize
minfit=minval(eval(:,1))
maxfit=maxval(eval(:,1))

!Normalizing Fitness
IF (maxfit==minfit) exit
DO i = 1,popsize
```

```
      eval(i,2)= (eval(i,1)-minfit)/(maxfit-minfit)
END DO

OPEN ( UNIT = 123, FILE = 'C:\Final Project\Project
Network\evaluations.txt', STATUS = 'UNKNOWN')
DO i = 1, popsize
WRITE(123,*) -1*eval(i,1)
END DO

!Finidng Sum, Average, Minimum and Maximum NormalizedFitness
sumnorm=sum(eval(:,2))
avgnorm=sumnorm/popsize
minnorm=minval(eval(:,2))
maxnorm=maxval(eval(:,2))
!Calculate Cumilative Normalized cost
eval(1,3)=eval(1,2)/sumnorm
DO i = 2,popsize
eval(i,3)=eval(i,2)/sumnorm+eval(i-1,3)
END DO

!WRITE(*,*)'.................................................'
! WRITE(*,*) 'Sum    = ', sumfit, sumnorm
! WRITE(*,*) 'Avg    = ', avgfit, avgnorm
! WRITE(*,*) 'Min    = ', minfit, minnorm
! WRITE(*,*) 'Max    = ', maxfit, maxnorm

! WRITE(*,*)'.................................................'

!!!!!!!!!!!!!!!!!
!                 Roullette Wheel Selection     !
!!!!!!!!!!!!!!!!!


IF (selection == 'R') THEN
ALLOCATE(m_rnd(1:popsize, 1))
CALL RANDOM_NUMBER(m_rnd)
randsel(:) = NINT(m_rnd(:,1) * 100000000.)/100000000.
DEALLOCATE(m_rnd)
parent=0


DO i = 1, popsize
       DO j=1, popsize
            IF (eval(j,3)>=randsel(i)) THEN
        EXIT
        END IF
      END DO
parent(i)=j
END DO
!!!!!!!!!!!!!!!!!
!                 Tournament Selection          !
!!!!!!!!!!!!!!!!!
ELSE IF (selection == "T") THEN
ALLOCATE(m_rnd(1:popsize, 2))
CALL RANDOM_NUMBER(m_rnd)
parent=0
m_rnd = NINT(m_rnd*(popsize-1)) + 1
DO i = 1, popsize
      IF ( m_rnd(i,1)>= m_rnd(i,2)) THEN
                  parent(i) = m_rnd(i,1)

      ELSE
                  parent(i) = m_rnd(i,2)
      END IF
```

```
END DO
DEALLOCATE(m_rnd)
ELSE
WRITE(*,*) 'Please select a Selection Method'
END IF
! WRITE(*,*)'.................................................'
!!!!!!!!!!!!!!!!!!!!!
!                  Selecting New Generation                   !
!!!!!!!!!!!!!!!!!!!!!

!WRITE(*,*)'========================='
!WRITE(*,*)'Selecting New Generation'
!WRITE(*,*)'========================='


offspring=ipop(parent,:)


!DO i = 1,popsize
!WRITE(*,40) parent(i), offspring(i,1:8)  , offspring(i,9:16),
offspring(i,17:24), offspring(i,25:32)
!END DO
!40 FORMAT(I3, '  ', 8I2,'  ', 8I2, '  ', 8I2, '  ', 8I2)
!!!!!!!!!!!!!!!!!!!!!
!                  Applying CrossOver                         !
!!!!!!!!!!!!!!!!!!!!!


!WRITE(*,*)'=================='
!WRITE(*,*)'Applying CrossOver'
!WRITE(*,*)'=================='



child= 0

DO i = 1, popsize, 2

CALL RANDOM_NUMBER(s_rnd)
check_cross=s_rnd

IF (cross_p >= check_cross ) THEN

      counter(igen) = counter(igen) + 2
      CALL RANDOM_NUMBER(s_rnd)
      cross=  NINT(s_rnd*(str*dv-1-2))+2
    DO j = 1, cross
      child(i, j)=offspring(i,j)
      child(i+1, j)=offspring(i+1,j)
    END DO

      DO j = cross+1, dv*str
      child(i,j) = offspring(i+1, j)
      child(i+1,j) = offspring(i, j)
      END DO
 else
      DO j = 1, dv*str
      child(i,j)=offspring(i,j)
      child(i+1,j)=offspring(i+1,j)
        END DO

END IF
END DO
```

```
!DO i = 1,popsize
!WRITE(*,50) child(i,1:8)   , child(i,9:16), child(i,17:24),
child(i,25:32)
!END DO
!50 FORMAT(8I2,'   ', 8I2, '   ', 8I2, '   ', 8I2)

!WRITE(*,*)'.................................................'
!!!!!!!!!!!!!!!!!
!                    Starting Mutation              !
!!!!!!!!!!!!!!!!!

!WRITE(*,*)'================='
!WRITE(*,*)'Applying Mutation'
!WRITE(*,*)'================='
DO i = 1, popsize*dv*str
      CALL RANDOM_NUMBER(s_rnd)
      check_mutate = s_rnd
      IF(check_mutate <= mutate_p ) THEN
                  counter(igen) = counter(igen) + 1
              CALL RANDOM_NUMBER(s_rnd)
                  mutate = NINT(s_rnd*(dv*str-1) +1)
                  CALL RANDOM_NUMBER(str_mutate)
                  i_mutate = NINT(str_mutate * (popsize-1)+1)

                  IF (child(i_mutate,mutate)==0) THEN
                 child(i_mutate,mutate)=1
            else
                   child(i_mutate,mutate)=0
              END IF
      END IF
END DO

!DO i = 1,popsize
!WRITE(*,60) child(i,1:8)   , child(i,9:16), child(i,17:24),
child(i,25:32)
!END DO
!60 FORMAT(8I2,'   ', 8I2, '   ', 8I2, '   ', 8I2)

!WRITE(*,*)'.................................................'
!!!!!!!!!!!!!!!!!
!                    Ellitism                       !
!!!!!!!!!!!!!!!!!
!WRITE(*,*)'================='
!WRITE(*,*)'Applying Ellitism'
!WRITE(*,*)'================='

best = eval(1,1)
bestindex=1

DO i = 1, popsize
      IF (eval(i,1)>= best) THEN
      best=eval(i,1)
      bestindex=i
END IF
END DO
counter(igen) = counter(igen) +1

IF (check_elitism == 'Y') THEN

      CALL RANDOM_NUMBER(s_rnd)
      elite=NINT(s_rnd*(popsize-1)+1)
      child(elite,:) = ipop(bestindex,:)

ELSE IF (check_elitism == 'N') THEN
```

```
! Do Nothing

ELSE

WRITE(*,*)'Please Make an Elitism Selection "Y" if yes, "N" if no.
Elitism cannot be left empty'

END IF


!DO i = 1,popsize
!WRITE(*,70) child(i,1:8)  , child(i,9:16), child(i,17:24),
child(i,25:32)
!END DO
!70 FORMAT(8I2,'  ', 8I2, '  ', 8I2, '  ', 8I2)
!WRITE(*,*)'...................................................'
!!!!!!!!!!!!!!!!!
!                  Preparing For Next Generation              !
!!!!!!!!!!!!!!!!!
ipop = child

!!!!!!!!!!!!!!!!!
!                  GoTo Next Generation            !
!!!!!!!!!!!!!!!!!
WRITE(*,*)'====================='
WRITE(*,*)'END OF GENERATION No: ', igen
WRITE(*,*)'====================='
WRITE(*,*)'............................................................
.................'
bestever(igen, 1 ) = eval(bestindex,1)
bestever(igen, 2 ) = avgfit

END DO
!!!!!!!!!!!!!!!!!
!                  Number of Function Evaluations                  !
!!!!!!!!!!!!!!!!!
fit = -fit
DO i = 1, popsize * generation
      DO j = 1, i
            IF ( fit(i) == fit(j)) THEN
            freq(i) = freq(i) + 1
            END IF
      END DO
WRITE(*,*) i
END DO
s = 0.
DO i = 1, generation
      DO j = (i-1)*popsize+1, i*popsize
            IF (freq(j) <> 1 )THEN
            s = s + 1.
            END IF
      END DO
      nevals(i) = popsize - s
      s = 0.
WRITE(*,*) i
END DO

!!!!!!!!!!!!!!!!!
!                  Final Resutls                !
!!!!!!!!!!!!!!!!!
WRITE(*,*)'====================='
WRITE(*,*)'THE FINAL RESULTS ARE '
WRITE(*,*)'====================='
best=eval(1,1)
```

```
bestindex=1
DO i = 1, popsize
IF (eval(i,1)>= best) THEN
      best = eval(i,1)
      bestindex = i
END IF
END DO

OPEN ( UNIT = 200, FILE = 'C:\Final Project\Project
Network\bestever.txt', STATUS = 'UNKNOWN')
cumcounter = 0
DO i = 1, igen-1
cumcounter = cumcounter + counter(i)
WRITE(200,*) i,  bestever(i, 1) , bestever(i, 2), counter(i),
cumcounter, INT(nevals(i))
END DO
WRITE(*,*)'...............................................'
WRITE(*,*) 'Best Index =', bestindex
OPEN ( UNIT = 300, FILE = 'C:\Final Project\Project Network\output.txt',
STATUS = 'UNKNOWN')
WRITE(300, 300) x(bestindex,  1:7 ), x(bestindex,  8:14)
WRITE(300, 300) x(bestindex, 15:21), x(bestindex, 22:28)
WRITE(300, 300) x(bestindex, 29:35), x(bestindex, 36:42)
call cpu_time(time_out)
write(300,*) time_out-time_in

300 FORMAT(7(F4.1, ' '), '*  ', 7(F4.1, ' '))

OPEN ( UNIT = 201, FILE = 'C:\Final Project\Project Network\sumary.txt',
STATUS = 'UNKNOWN')
call cpu_time(time_out)
WRITE(201,*) cross_p, mutate_p, INT(icase/100), bestever(igen-1,1),
time_out-time_in, INT(SUM(nevals))

WRITE(*,*) 'Closing..............................'

DEALLOCATE(ipop)
DEALLOCATE(offspring)
DEALLOCATE(child)
DEALLOCATE(var)
DEALLOCATE(x)
DEALLOCATE(eval)
DEALLOCATE(randsel)
DEALLOCATE(parent)
DEALLOCATE(bestever)
DEALLOCATE(counter)

write(*,*) time_out

STOP

END PROGRAM
```