NORTHWESTERN UNIVERSITY

ASSESSING THE IMPACT OF CONNECTED VEHICLES AT FREEWAY, ARTERIAL, AND PATH LEVEL: CHARACTERIZATION, MODELING AND ACTIVE MANAGEMENT

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

For the degree

DOCTOR OF PHILOSOPHY

Field of Civil and Environmental Engineering

By

Archak Mittal

EVANSTON, ILLINOIS

December 2018

ABSTRACT

ASSESSING THE IMPACT OF CONNECTED VEHICLES AT FREEWAY, ARTERIAL, AND PATH LEVEL: CHARACTERIZATION, MODELING AND ACTIVE MANAGEMENT

Archak Mittal

The term Connected Vehicle (CV) is broadly used to identify any 'smart vehicle' with wireless connectivity to the roadside infrastructure and other vehicles. CVs with automation capabilities are called connected automated vehicles (CAVs). With real-time communication and data transmission capability, CAVs have the potential to improve the transportation system's traffic flow, reliability, and safety. However, due to lack of data on CAVs their potential has not been fully realized. Consequently, this dissertation explores the potential of CAV technology in formulating 1) facility type-customized active management and control strategies, 2) assessment techniques, and 3) analytical methods with the purpose of enhancing existing transportation network's operational capabilities.

To explore the potential of CAVs, the dissertation is divided into three components. In the first component, the dissertation focuses on freeways and highways. At this level, study first puts forward a computationally efficient framework to model CAVs using existing simulation tools and relevant data. Using this framework, traffic flow conditions and travel time reliability were studied. Results of this framework show that CAVs can improve traffic flow conditions while increasing the reliability of the system under different demand levels and operational conditions including inclement weather condition.

In the second component, the dissertation focuses on the arterial roads in a transportation network. In this component, microscopic traffic flow models were utilized to emulate CAVs on the arterial road and formulate two advanced traffic signal control strategies. Two advanced traffic signal control logics are developed, via V2I and V2V communication. These signals provide synchronized traffic flow on major corridors while keeping the logic decentralized. CAVs compute their travel time delay accumulated on a route. This accumulated delay forms basis of decentralized but coordinated traffic signal strategies. Numerical experiments show that the decentralized but coordinated traffic signal strategies outperform state of the art practices.

In the third component, the analysis is conducted at the user-defined path level. A user-defined path can consist of both arterial and freeways. Thus, the study of user-defined path opens an interesting avenue to analyze the combination of arterial and freeways and how they interact with each other. The analysis in this component formulates an innovative network partitioning concept based on the average path-level fundamental relationships among the traffic stream variables. Correspondingly, time-of-day control and management strategies can be tailored to suit specific paths' operational characteristics. Furthermore, to adequately compute link travel time correlations and accurately determine path-based travel time distributions an analytical model was designed. Analytical form of path travel time variance was devised to correctly capture the spatiotemporal covariance of link travel times. Travel time distributions along the paths defined by users were estimated through solving a convoluting integral of correlated link travel times. Numerical experiments show that the model accurately estimates the travel time distribution along paths.

The developed simulation techniques, control strategies, and assessment methods should be used to enhance model calibration and validation; enrich system's performance, performance

evaluation; and provide a sound basis for making routing decisions taking quantifiable risk estimates into account.

ACKNOWLEDGMENTS

I would like to thank my advisor, Prof. Hani Mahmassani, for his time, effort, support, guidance, and encouragement. He has been a great supervisor, teacher, and mentor throughout my doctoral years. It has been a true privilege for me to work with him. I cannot thank him enough for the trust and confidence that he has shown in my work. He is the foundation of my professional life that I will pursue henceforth.

I am thankful to the members of my dissertation committee, Prof. David Morton, Prof. Yu (Marco) Nie, Prof. Ying Chen, for giving me invaluable comments and suggestions and for sharing their knowledge and experience throughout the process of this dissertation. I am also grateful to Prof. Amanda Stathopoulos, Prof. Joseph Schofer and Prof. Pablo Durango-Cohen for their teaching and inspiration that had helped me to deepen and broaden my knowledge in the field of transportation engineering.

I have been fortunate to work with such a talented group of people at the Northwestern University Transportation Center who have supported me throughout this process both as colleagues and as friends. Zihan, you were my first peer mentor after I joined here. I learned about conducting research and handling stressful situations from you. I am thankful to my Habibi group of Lama, Amr and I. Lama, you have been a true friend, the best coffee partner, great guide. Amr, you have been a great support and squash buddy. Habibis have been a great support for me to complete the journey during my Ph.D. Mike, thank you for being my American mom, sharing your wisdom, tolerating my irritating jokes and puns. Marija, I wish there were words to thank you enough for being there, literally, always. You made this journey easier for me. I could not have done it without you all. With the doctoral degree, I also earned a friendship for a lifetime.

I would also like to thank Small Lama, Monika, Eunhye (InHey), Xiang (Alex), Alec, Aymeric, Dr. Ömer Verbas, Dr. Alireza Talebpour, Dr. Ali Zockaie, and Dr. Hooram Halat. Several staff members of the Northwestern University Transportation Center and the Department of Civil and Environmental Engineering greatly enriched my experience as a graduate student. I am particularly grateful to Diana Marek, Hillary Bean, Bret Johnson, Joan Pinnell, and Cynthia Ross for their helpfulness, kindness, and friendliness.

Last, but certainly not least, I would like to express my deepest gratitude to my parents for their unconditional love and support. Their endless belief and confidence in me gave me the power to accomplish this work. They always put my needs first. I owe everything to them. I dedicate this dissertation to them.

ABBREVIATIONS

AFD	Arterial Fundamental Diagram
ARIMA	Autoregressive Integrated Moving Average
ASLV	Adjusted Spatial Longitudinal Variation
CACC	Cooperative Adaptive Cruise Control
CAR	Connected, Automated and Regular
CAV	Connected Automated Vehicles
CTR	Cumulative Travel-time Responsive
CV	Connected Vehicles
DSRC	Dedicated Short-Range Communications
DTA	Dynamic Traffic Assignment
DTW	Dynamic Time Warping
GPS	Global Positioning System
ID	Identity
IDM	Intelligent Driver Model
IGA	IntelliGreen Algorithm
ITS	Intelligent Transportation Systems
KM	K-Means
MFD	Macroscopic Fundamental Diagram
MOE	Measure Of Effectiveness
MPR	Market Penetration Rate
MWM	Maximum Weight Matching
NEMA	National Electrical Manufacturers Association

NFD	Network Fundamental Diagram
NGSIM	Next Generation Simulation
OD	Origin–Destination
PAMSCOD	Platoon-based Arterial Multi-modal Signal Control with Online Data
PFD	Path Fundamental Diagrams
PMSA	Predictive Microscopic Simulation Algorithm
P-SCO	Prediction Signal Control Optimization
RBC	Ring Barrier Controller
SchIC	Schedule-driven Intersection Control
STARIMA	Space-Time Autoregressive Integrated Moving Average
TDTC	Two-Dimensional Time Correlation
TMC	Traffic Management Center
V2I	Vehicle to Infrastructure
V2V	Vehicle to Vehicle
VCA	Vehicle Clustering Algorithm
WCSS	Within-Cluster Sum of Squares

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. INTRODUCTION

The pace of transportation innovation is accelerating towards safer, smarter and greener systems. Car manufacturers and researchers are designing and producing vehicles which are able to generate information regarding the vehicle's performance, state, and trajectory – time-stamped position, heading, speed, routing, and driving style. Previous studies have established that trajectory-based measurements are more accurate than those based on information aggregated over road segments. Current transportation research and practice will benefit from trajectory-based studies as they provide more insightful and accurate measures for the state of the system. These aggregated measurements, such as travel time delay or variation of travel time, lose their underlying distribution and the extreme values get smoothened out. Connected Vehicles (CV) are equipped with the capability to communicate amongst themselves and with the traffic control infrastructure. Using the communication capability, CVs can share their data on how, when, and where they travel. Such data can help vehicle drivers to optimize their route and planners to predict the state of the system and decide the control strategy parameters. This dissertation explores the potential of CVs to make the transportation system more efficient, responsive, agile, reliable, and safe.

CVs can provide data about their departure time, origin and destination of trip, route, condition, and features of the vehicle. CVs can also provide their trajectories, which consist of time-stamped GPS positions. The produced data is transmitted to nearby traffic control infrastructure, a.k.a. Vehicle to Infrastructure (V2I) communication, or can also be shared with vehicles traveling in the vicinity, a.k.a. Vehicle to Vehicle (V2V) communication, using dedicated short-range communication (DSRC) protocol. In simple terms, DSRC can be seen as a short-range Wi-Fi connection established by the vehicles. Hence, CVs can transmit the location and progression-

related information in a high-resolution format and at a faster rate than any other existing technology. Such capabilities of CVs can be utilized to improve the traffic flow condition and to increase the reliability of the transportation system. Some CVs may additionally have automated vehicle (AV) features. AVs are equipped with various types of detectors to sense the vicinity of a vehicle. Based on the information from the vicinity, vehicles make decisions to maneuver. AVs have practically zero reaction time, especially when receiving information from nearby CVs, allowing them to react quickly to emergency events. CVs with the features of AVs can be envisioned as the ultimate mode of automobile transportation. Connected automated vehicles (CAVs) have the potential to improve the transportation system's traffic flow, reliability, and safety. CAVs are proactive, cooperative, and well-informed of their surroundings, thus, pave the way for supporting various applications. Such capabilities of CAVs motivate research to explore their potential to the greatest extent possible. The research focus, in the past, has been on the three principal areas: roadway safety, the mobility of vehicles, and eco-friendly standards. Similarly, in this dissertation, CAV enabled applications are being developed to promote the same transportation standards.

Due to lack of information on each vehicle on roads, researchers estimate the traffic flow variables using measurements from road sensors. However, measurements based on individual vehicles' trajectory information are more accurate than those based on aggregated information. With CAVs, researchers can obtain information on individual vehicles, and as a result information on traffic stream variables can be directly obtained. Additionally, the individual trajectory-based information offers high-resolution and precise information on the transportation system's state to predict

events, such as a traffic breakdown, or the occurrence of gridlock, and devise control measures to mitigate them, enabling more comprehensive traffic state characterization, analysis, and control.

While studying CAVs' potential, it is important to understand the environment in which they operate i.e. a transportation network. A transportation network is a grid of roads spread throughout an area. These roads intersect each other at nodes. A road segment connecting two nodes is a link. These links can belong to various transportation facilities: arterial roads, urban streets, freeway, highway, ramps, and alleys. Traffic flow patterns differ by link type, even when the same drivers traverse diverse types of facilities. Based on the type of traffic flow, link types can broadly be categorized into two groups as uninterrupted and interrupted. Uninterrupted traffic flow facilities, such as freeways and highways, are regulated by vehicle-to-vehicle interactions and their interactions with the roadway. Once on an uninterrupted facility, vehicle flow is not interrupted by the infrastructure of the facility. Interrupted flow facilities, such as arterial roads, are regulated by external means, most commonly traffic signals and stop signs. This dissertation developed two advanced traffic signal control strategies to design a smart and environment-friendly arterial corridor. For freeway facilities, platforms that enable faster movement of vehicle platoons, such as speed harmonization can be designed. A user generally takes a trip on paths comprised of different types of links. They are not constrained to one type of links. For example, a user may start on an urban street, move to a freeway and then again take an urban street to arrive at its destination. In this case, a good service just on the freeway segment does not guarantee a good overall experience for the user. A travel time delay incurred on urban streets will count towards the experience. Accordingly, analysis at a user-defined path level is essential to understand a user's

experience and identify and address the issues accordingly. A path-level study requires rich, high-resolution data that could be provided by CAVs.

This dissertation proposes to answer the following fundamental research questions:

- How can connected vehicle data and technologies be used to support offline and online performance-based management of facilities and transportation systems overall?
- How much can the transportation system's (or its elements') mobility, efficiency, and environmental standards be enhanced in a meaningful manner, when only a fraction of the traffic stream is connected?
- Can connectivity help isolate the underlying causes of system inefficiencies and facilitate the development of innovative analytical methods to describe and address these issues?
- How can connected vehicle-generated traffic data be utilized for traffic state characterization and reactive and predictive analytics with respect to different operational conditions?

These questions are addressed through several applications using a variety of simulation tools at different levels of spatial and temporal resolution, as described in the remainder of this chapter.

1.1. Motivation and objectives

CAVs can provide high-resolution trajectory data. Communication capability of CAVs enables a faster transaction of information. With CAVs, the state of a system can be assessed in real time, promptly and accurately. The opportunities around connected vehicle data are numerous. However, an extremely low market penetration rate (MPR) of CAVs on existing road facilities has

limited the data available on CAVs to date. Simulation-based studies are a widespread practice among researchers when data is scarce and when developing or evaluating a modern technology or new concept. Simulation tools and the models within them need calibration and validation, against the ground truth. If simulation outputs are within a tolerance of ground truth (typically within 10% tolerance), the tools/models are considered reliable and consequently results credible. Along the same line, this study first puts forward a computationally efficient framework to model CAVs using already existing simulation tools. The framework is tool independent and transferable to any simulation setting as per the analyst's choice. To demonstrate its feasibility, in-house dynamic traffic assignment (DTA) tools and a commercially available microsimulation tool were used. Different MPRs of CAVs and operational conditions were examined. The conducted analysis disaggregates a transportation network in multiple ways to account for different demand levels, operational conditions (such as weather), and types of facility and levels of analysis. This effort models a connected environment and explores its potential as both a data source and an application platform. Even though the models are based on simulation-generated traffic data, they are readily applicable to real-world vehicle trajectories. Applicability of the models to the real-world vehicle trajectory data is demonstrated in Chapter 9 of the dissertation.

Current fixed sensor data as currently used in practice provide aggregated information on a facility's performance, where extreme measurements of the collected information are lost. Transmission and processing of the data to optimize control measures experience a lag in time. As mentioned earlier, these control measures differ by the transportation facilities. Further lag is encountered when the optimized control measure is conveyed back to the drivers. CAVs can help manage the system better by providing real-time and high-resolution data and immediate feedback

message(s). Accordingly, this research examines the operational benefits of connectivity through facility type-customized active management strategies. Therefore, in this study, connected vehicles attributes and functionalities are investigated at four levels: 1) aggregate level, where the entire transportation network is treated as one entity; 2) freeways and highways level, where information on links of these types is aggregated together, 3) arterials roads, where information on arterial type links is aggregated together, and 4) user-defined path level, where a path may comprise links of any type, such as freeway, highways and arterials, and are consecutive. For the last level, no control strategy was developed as it is addressed in the second and third level of aggregation. At the path level, the study formulates a novel clustering method for user-defined paths, and study the fundamental diagrams of the path clusters.

Analyzing rich trajectory data reveals meaningful connections, trends, and patterns that can help provide a better driver experience and improve transportation systems efficiency, quality and reliability. For this reason, the study is conducted on several aggregation levels - microscopic, mesoscopic, and macroscopic level - to identify and quantify the potential effect of connectivity on vehicle interactions, driving behavior, and traffic stream variables key relationships and characterize the macroscopic nature of the system as a whole.

Quantification of link travel time correlations is a concept that has not yet been applied to any extent in actual practice. This concept makes use of the rich, high-resolution data offered by vehicle trajectories. One of the reasons for this concept to be at a nascent stage is the unavailability of relevant data. To address this gap, an analytical framework is developed to adequately compute link travel time correlations with the aim of accurately determining path-based travel time

distributions. The findings are intended to support an evaluation effort that will inform a broader cost-benefit assessment of connected vehicle concepts and technologies.

1.2. Objectives and Contributions

The primary objective of this dissertation is to explore the potential of CAV technology to improve the performance of the transportation system. The objectives of this study are decomposed into three major components:

1. Modeling and assessing the impact of connected vehicles at a large network level
2. Control strategies under a connected environment with mixed vehicular traffic
3. Analytical assessment techniques for user-defined path-based analysis

These are three independent components focusing on different aspects of transportation modeling and engineering in a mixed connected vehicle environment.

1.2.1. Modeling and assessing the impact of connected vehicles at a large network level.

As mentioned earlier, due to low MPR of CAVs, relevant data and infrastructure are not available to the researcher. Therefore, simulation technique in a testbed context constitute the initial focus of this dissertation. Here, CAVs as a data source and the latter's impact on traffic operations is evaluated. The framework models communication between transportation system's infrastructure and CAVs through a hybrid of microscopic and mesoscopic simulation models. Specifically, *drivers' behavior is emulated at the microscopic level and transferred to a mesoscopic simulator*

for network level assessment. The framework is computationally efficient at large scale, independent of a simulation tool, and transferable to any simulation setting as per analyst's choice. The feasibility of the framework has been demonstrated under different market penetration rates and operational conditions, using in-house DTA tools as well as commercially available microsimulation tool. To analyze the effect these vehicles have on the surrounding traffic, vehicle trajectory-based fundamental diagrams of traffic flow at the network level, known as network-wide fundamental diagrams (NFDs) or macroscopic fundamental diagrams (MFDs), and travel time reliability measures were used.

1.2.2. Control strategies under a connected environment with mixed vehicular traffic.

To study the impact of new technologies on signalized urban corridors, an adaptive, prediction-based traffic control strategy is developed to manage mixed traffic environments while optimizing vehicle trajectories to meet an eco-friendly objective. To further examine the communication-enabled signal control, a real-time platoon self-identification traffic signal control strategy based on the V2X communication technology is developed and tested. These advanced intersection-level control strategies also help to maintain coordination on a corridor. To achieve coordination, *CAVs, in this study, are assumed to compute and transmit their accumulated delay* along a corridor or a facility, thus enabling corridor traffic flow synchronization when incurred delay justifies such actions.

1.2.3. Analytical assessment techniques for user-defined path-based analysis.

Users' experience in a transportation system that comprises multiple facilities and sometimes multiple modes of travel is captured through their trajectories. It is important to study the paths that are taken frequently by the users. Popular paths can be identified using trajectories of the vehicles and clustered to find distinguishing patterns among them. Thus, user-centric measures of system performance are sought to translate objective measurements of attributes such as travel time and delay along links to overall time and delay measures at path level. Accordingly, *an analytical framework is developed to adequately compute link travel time correlations* with the aim of accurately determining path-based travel time distributions. Furthermore, the study proves that fundamental diagrams exists at path level and formulates an innovative network partitioning concept based on the path fundamental diagrams. The findings are intended to support an evaluation effort that will inform a broader cost-benefit assessment of connected vehicle concepts and technologies.

1.3. Organization of the Dissertation

This section describes the organization of the dissertation. Figure 1-1 represents an overall analysis framework, comprising a transportation system with its major components as well as external factors that affect its performance.

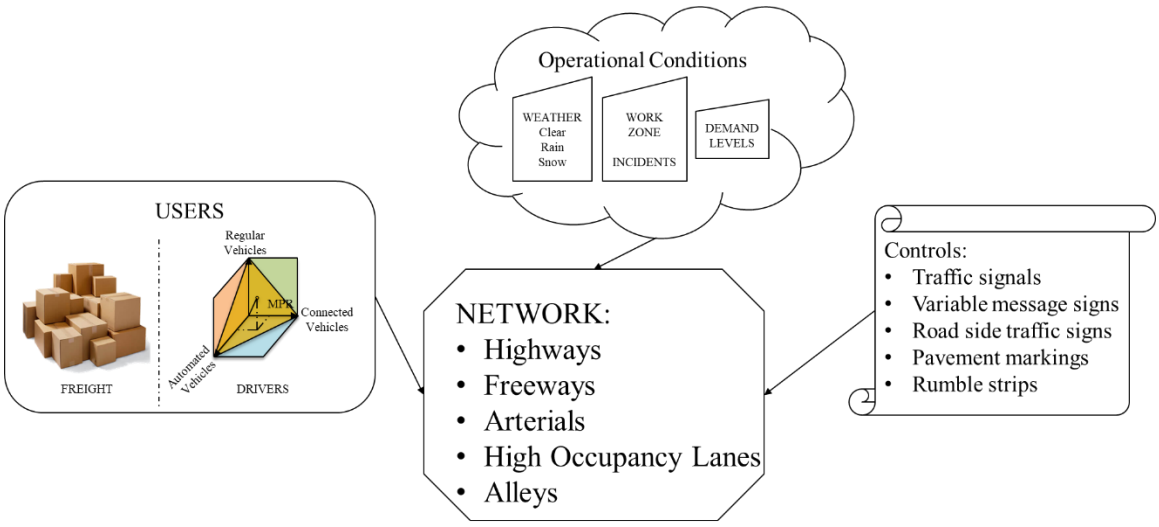


Figure 1-1 Overall Framework.

The focus of this dissertation is on the surface (private) vehicular traffic. The elementary infrastructure of this transportation system consists of facilities of different types, which will be referred to as the "network" henceforth. The network is used to move the two primary types of users, namely drivers (and their passengers) and freight. Here the author's focus is on the drivers. This dissertation starts with a thorough review of the existing studies in **Chapter 2: LITERATURE REVIEW**, with emphasis on modeling, performance evaluation, operational benefits as well as new analytical methods and concepts not yet discussed in the previous studies. The focus of the literature review is on operational conditions, facilities and applications related to CV- enabled environments. Automated and connected vehicles are becoming a reality and soon will make a sizable proportion of vehicular traffic. Currently, CAV technology is not yet market-ready. Consequently, there is not much information available on the potential impact of CAV technology on the system and its application-based benefits in a variety of traffic conditions. Researchers are circumventing this limitation by conducting simulation-based studies.

Accordingly, **Chapter 3: INTEGRATED SIMULATION FRAMEWORK DEVELOPMENT** puts forward a framework that integrates microscopic and mesoscopic modeling of connectivity in a large-scale traffic environment. The framework preserves microscopic details of drivers' behavior of different vehicle types operating in various traffic conditions. Using microscopically calibrated fundamental diagrams a large-scale network is then macroscopically simulated. Next, **Chapter 4: IMPACT OF CONNECTED VEHICLES** estimates the impact of connected vehicles on overall traffic operations by applying the framework presented in the previous chapter. Connected vehicles are introduced along freeways and highways. To analyze the effect of CVs on the surrounding traffic, fundamental diagrams of traffic flow and travel time reliability measures were studied. Measures of effectiveness or operational performance indicators were computed based on individual vehicle trajectories. Various MPRs of connected vehicles are studied to study an incremental benefit of the technology.

Chapter 5: ADVANCED SIGNAL CONTROL STRATEGIES introduces two signal control strategies for arterial roads. The first strategy is based on predictions of the future state of the system and the second strategy is based on self- identified vehicle platoons. The two signal control strategies' introduction in this chapter serves as a precursor for the study on the impact of new technologies along signalized arterial corridors, in **Chapter 6: PREDICTION-BASED ADAPTIVE SIGNAL CONTROL**. In Chapter 6, an adaptive, prediction-based, traffic control strategy was developed to manage mixed traffic environments while optimizing CAVs vehicle trajectories to meet an eco-friendly objective. To further examine the connectivity-enabled signal control aspect along the same direction, real-time applications are considered in **Chapter 7: REAL-TIME PLATOON SELF-IDENTIFYING ADAPTIVE SIGNAL CONTROL**. A real-

time, platoon self-identification, control strategy based on the V2I communication technology is developed and tested, and significant findings are presented in Chapter 7.

A user is not confined to use any particular type of roadway while commuting from one point to the other. Previous studies are based on fundamental diagrams of either links or the whole network. When conducting a study at user-defined path, a question arises, do fundamental diagrams exist at path levels? Hence, a user-defined path level study is undertaken to establish path-based fundamental relationships between traffic parameters. Accordingly, in **Chapter 8: PATH FUNDAMENTAL DIAGRAMS,** a new clustering technique is designed to disaggregate a network by paths taken by the users. Popular paths were identified using trajectories of the vehicles and were clustered to find distinctive patterns among them. A new theoretical basis for clustering paths and a novel clustering technique is formulated. With the identified path clusters, fundamental diagrams were studied to characterize the clusters. In **Chapter 9: TRAVEL TIME DISTRIBUTION ALONG USER-DEFINED PATHS**: an analytical model to describe travel time distributions at a user-defined path level is developed utilizing vehicle trajectories. The model also computes the spatiotemporal covariance of link travel times.

CHAPTER 2. LITERATURE REVIEW

This chapter summarizes the most relevant past research on the topics addressed in the dissertation. The literature review section is divided into several sub-chapters. The first focuses on studies modeling connected vehicle communication and interactions in micro/meso simulation environments. The emphasis is on understanding the effects of connectivity on traffic flow dynamics at the network level. The second summarizes past research on selected connected vehicle applications – i.e., dynamic speed limits – focusing on quantifying potential operational benefits. The third subchapter, unlike previous two, focuses on arterial-level applications of connected vehicle technology, with reference to individual vehicle (trajectory) information and its worthiness in developing advanced signal control strategies. Finally, relevant literature on techniques of identifying relevant paths in a transportation system, composed of the freeway and arterial links, are considered. Its utilization for connectivity-enabled network-level active traffic management schemes is discussed.

2.1. Connected vehicles in the network

There is a growing need for mobility and safety in transportation systems and cities are undergoing constant changes to accommodate this need. However, cities face limited resource availability to expand the current road networks. To address the challenge, cities are incorporating modern technologies into their current operational practices. Connected Vehicles technology, as one of the latest technologies in surface transportation, provides the opportunity to create a connected network of vehicles and infrastructure. In this network, individual vehicles can communicate with

each other through Vehicle-to-Vehicle (V2V) communications. Moreover, individual vehicles are connected to infrastructure and Traffic Management Center (TMC) through Vehicle-to-Infrastructure (V2I) communications. The real-time information provided by this technology can improve drivers' operational, tactical, and strategic decisions; thus, it can improve mobility and safety and reduce emissions and energy consumption in transportation systems.

Despite a considerable number of studies that investigated the effects of different applications of Connected Vehicles technology (e.g. speed harmonization, cooperative adaptive cruise control, queue warning, and transit signal priority) on mobility and safety (*2-7*), the effects of this technology and its application at the network level has not been investigated to the same extent. Understanding the effects of connectivity on traffic flow dynamics at the network level is critical to assess the effects of this technology on surface transportation systems. Many previous studies observed a well-defined network-level relationship between flow and density (*8-13*). Aside from the pioneer works of Smeed (*14*), Thomson (*15*), Wardrop (*16*), and Godfrey (*17*) on investigating this relationship, Herman and Prigogine (*18*) showed that their two-fluid model could characterize the relationship between the fraction of moving vehicles and average velocity. Following the finding, several other studies investigated traffic flow dynamics at the network level (see Chang and Herman (*19*), Mahmassani et al. (*10*), and Williams et al. (*11*) for some examples).

Recent advances in data collection and availability of large-scale, high-quality data create an opportunity to revisit the network-wide traffic flow relationships. Accordingly, Daganzo (*20*) further formalized the flow-density (or flow-occupancy) relationship at the network level, referring to it as a Macroscopic Fundamental Diagram (MFD). Following this study, several other studies investigate this relationship to characterize the observed patterns at the network level (see

Geroliminis and Daganzo (*9*), Ji et al. (*21*), Mazloumian et al. (*22*), Saberi and Mahmassani (*13*), Saberi and Mahmassani (*12*), Gayah et al. (*23*) for a few examples). There are also several studies that proposed different methodologies to estimate/characterize the network-wide flow-density relationship based on vehicle trajectory data. Saberi et al. (*1*) proposed a method to characterize this relationship using three-dimensional vehicular trajectory data at the network level. Nagel and Gayah (*24*) proposed a method to estimate this relationship using trajectories from probe vehicles at different market penetration rates (MPRs). These studies suggest that similar methodologies can be utilized to capture the effects of connectivity on surface transportation systems at the network level. Connected Vehicles technology is expected to have a significant effect on the characteristics of the network-wide flow-density relationship. Less significant hysteresis loops, better recovery, and less gridlock is expected since this technology can improve capacity and increase traffic flow stability (*4*).

2.2. Adaptive traffic signal strategies for urban roads under a connected environment

Traffic signals are the most common control strategy on arterial roads. Signal settings that are responsive to the changing traffic conditions can alleviate traffic congestion and consequently reduce associated delays. Actuated signal control strategies address the drawbacks of pre-timed signals by being responsive to the changes in the demand. Furthermore, adaptive signals are based on continuous monitoring of arterial traffic conditions and queuing at intersections as well as the dynamic adjustment of the signal timing to optimize one or more operational objectives (minimize delays, maximize throughput, etc.) (*25*).

Recently, signal control studies using wireless communication or more specifically CV technology came to be widely popular. As CVs collected data provide a much more complete picture of the arterial/intersection traffic states, opportunities to harness these for control purposes became evident (*26*).

Consequently, CV-based signal control strategies rely on more accurate detection and in the case of rolling horizon approaches (*27-30*) offer a more reliable prediction.

Goodall et al. (*29*) proposed a predictive microscopic simulation algorithm (PMSA) for signal control, which used data from connected vehicles including positions, headings, and speeds and imported them to a microscopic simulation model to predict the future traffic conditions. Other researchers detected the presence of platoons in advance and optimized signal control parameters to accommodate the incoming vehicles at a downstream intersection (*31; 32*).

In addition, unlike conventional adaptive signal control systems that are restricted by the fixed location sensors, queue spillback issues under over-saturated traffic conditions can be addressed using CV technology (*33; 34*).

Despite adaptive strategies' successful implementation, their performance relies on the continuous and reliable operation of the detectors. The advent of V2V and V2I communication through DSRC is envisioned as a solution to the failure of detectors. Moreover, DSRC provides more information related to individual vehicles' travel. Real-time position, and under the specific concept of operations, desired route, destination, etc. are readily available. A comprehensive review of adaptive signal control strategies in a connected vehicle environment is presented by Jing et al. (*35*).

Past research distinguishes adaptive signal strategies based on dedicated short-range communications (DSRC) aimed at minimizing delay or travel time (*36-42*), queue length (*27; 43-48*), waiting time (*49-51*), (only implicitly) pollutant emissions (*49; 52*), number of stops (*53; 54*), fuel consumption (as one objective among multiple) (*38; 55*), or maximizing throughput (*47; 56; 57*) using data acquired through Connected vehicles (CVs), V2X communication, or vehicular ad hoc networks. However, relevant literature was found to be rather sparse in addressing the performance of advanced signal control strategies in mixed traffic conditions, at various MPRs of different vehicle types.

Real-time control applications based on platoon recognition methods are relatively rare. Computational complexity as well limits its applicability in a variety of traffic conditions. In recent years, however, platoons of connected vehicles have been the basis of several control techniques. Real-time connected vehicles' positions and speeds are essential information, which determines their arrival times when identifying or segmenting a platoon for traffic control purposes.

Gradinescu et al. (*42*) were one of the first to propose a phase sequence/duration optimization, to minimize control delay and queue length using car-to-car and car-to-controller communication. The authors utilized Webster's formula (*58*) to derive the amount of green per phase.

Wunderlich et al. (*59*) proposed a queue size based maximum weight matching (MWM) control framework, which became the benchmark for many other researchers when developing platoon-based strategies. The weights reflect the service urgency of each queue. The algorithm evaluates the size and weight of each queue and schedules phases to maximize throughput. Its flexible phasing setup provides superior performance regarding average vehicle delay compared to a sequential dual-ring phase scheme.

Accounting for minor shortcomings of the earlier approach by Wunderlich et al. (*59*), other models were formulated to account for variation in queue discharge rates. Longer queues discharged rates may be lower than those of shorter queues and shared lanes traffic mixture (straight and right/left-turning vehicles) impact on queue clearing times is carefully inspected (*60*). By leveraging turn information and vehicle lane positions, the control method selects the next best phase and decides its duration for overall (intersection) throughput maximization, while promoting "fairness": allocating more green time to those approaches with higher passing rates, and "occasionally" assigning right of way to lower ones for fairness provision.

Please note that fairness, in the context of this study, signifies priority with respect to accumulated (total) delay of connected and regular vehicles.

By utilizing the properties and relative positions of vehicles, attempts have been made to design vehicle scheduling-based control strategies (*28; 61; 62*). These methods incorporated look ahead horizons to either anticipate queues or platoons of vehicles. Similarly, given high-resolution vehicle trajectory information, advanced priority and pre-emption techniques were investigated (*31; 52*). Many subsequent approaches used these features in designing connectivity-enabled control algorithms since timely and accurate information is made available to measure what was once only estimated: queues and platoons of vehicles.

As part of a broader initiative to design advanced control strategies for connected vehicle environments, Smith et al.(*33*), designed multiple control strategies using intelligent transportation systems' data as the primary data source: queue identification and monitoring, vehicle clustering, and rolling horizon approach to optimize offsets and splits at signalized intersections. A vehicle clustering algorithm (VCA) was proposed, to find a suitable gap among approaching vehicles for

each phase's green end. The designed algorithm operates in three stages. The first calculates cumulative waiting times on each red-indication movement, second ensures gap out occurs as soon as the last vehicle has cleared the approach and the third, based on vehicles' distances, determines which pseudo-platoon is the closest to the intersection, yet farther than a certain threshold distance. Accordingly, appropriate green-extension times are computed. After the extension, the time has elapsed, or the maximum time is reached, the right of way is given to the red-movement that had the highest cumulative waiting time.

Extending their previous work (*33*) on fully connected control algorithms, Datesh et al. (*63*) developed an algorithm which determined the end-of-phase time by identifying a sharp decrease in vehicle density, calculated based on vehicle's location and distance to the intersection. Their IntelliGreen Algorithm (IGA) uses k-means clustering to determine the optimal point in time to terminate a green phase. A natural break in the time-to-intersection distribution of the vehicles approaching the green signal partitions the vehicles into two clusters: green and red. The largest time-to-intersection value in the green cluster is set as the remaining green phase time while the "red cluster" is stopped. However, traffic flow was divided into at most two platoons, regardless of the actual arrival pattern. Similarly, a Schedule-driven Intersection Control strategy (SchIC) was designed to efficiently produce (near) optimal solutions in real time (*28*). SchIC reduces the search space by exploiting queue size and temporal arrival distribution in the prediction horizon. Outside of the conventional approaches to traffic control, Venkatanarayana et al. (*34*) devised an algorithm to monitor the queue length at an intersection in real-time and adjust offsets and splits of the upstream intersection in response to queue spillback. The procedure adjusts the upstream movement's green phase either by delaying or shortening it. However, the algorithm was proved

to be active by reducing total delay only in a very simple network with 2 one-way streets intersections. The phase with the highest combined travel time was set to be the next in Lee et al.'s (*37*) cumulative travel-time responsive (CTR) real-time connected vehicle-based signal control algorithm for isolated intersections.

Assuming advanced communication between vehicles and traffic controllers, He et al. (*32*) formulated an offline arterial traffic signal optimization framework for multiple travel modes named Platoon-based Arterial Multi-modal Signal Control with Online Data (PAMSCOD). A headway-based platoon recognition algorithm categorizes individual vehicle requests and clusters them into platoons by priority level and phase. The procedure assumes first come first serve rule and aggregating vehicles into platoons which request priority to address the issue of computational complexity. Another feature of PAMSCOD is its ability to control the discharge rate from upstream intersections to avoid the de-facto red since real-time information regarding queue length and size are available. Under the same V2I framework, the authors structured a simplified formulation and a heuristic algorithm for real-time applications (*31*). Multiple priority requests from different modes are explicitly accommodated while simultaneously considering virtual priority requests for coordination and vehicle actuation. When coordination is broken, a penalty will reflect it in the objective function calculation.

As part of the initiative to design advanced control concepts for connected environments, Skabardonis et al. (*64*), developed and tested through simulation, a queue spillback avoidance strategy to improve mobility based on CV data. The method was formulated as a platoon-based control method. It comprised three distinct strategies: green extension, phase termination, and double cycling. Applying the most effective one depending on the associated total delay predicted.

The literature referenced, throughout, has one aspect in common: the most critical parameter in determining the effectiveness of control algorithms was the market penetration rate of connected vehicles. Most of the previous studies referred to offline (optimization) control strategies with 100% penetration rate of connected vehicles. Feng et al. (*30*) developed an adaptive signal control for CV-enabled isolated intersections where a two-level optimization problem that minimizes total vehicle delay and the queueing length is solved in real-time. Later the authors extended this methodology to one that integrates coordination with adaptive signal control in a connected vehicle environment (*65; 66*).

Alternative platoon-based methods pre-establish a platoon's size, disregarding that traffic demand can be disproportionate to different approaches or that individual vehicles might incur extremely high delays. Pandit et al. (*67*) instead of identifying emerging breaks in traffic flow, divide incoming traffic flow into partitions, which can be treated as equal-sized jobs in a processor job scheduling problem. The platooning represents an exhaustive search over all the platoon configurations to determine the platoon combination that minimizes the difference between the maximum and minimum green times. Then, the oldest arrival/job first algorithm generates a conflict-free job schedule with respect to minimizing delay. The control scheme was shown to perform as well as vehicle-actuated type in heavy demand. Very similar work has been done by Kokiladevi and Kumar (*2*) who define a threshold-based partitioning algorithm, which partitions vehicles into platoons based on their temporal distance from the intersection. Earliest arrival first algorithm, then, schedules the right of way of platoons of vehicles in a safe conflict-free manner. Cumulative platoon headway defines the actual phase duration. Simulation-tested, the method outperformed pre-timed and Webster's (*42*) controller logic, yet produced similar average vehicle

delay as a vehicle-actuated control. Under lighter traffic, the earliest arrival first algorithm can dynamically skip through phases and minimize the delay of vehicles whenever there is a gap in the traffic. However, platoon identification thresholds seem not to be actual demand-driven but subjective.

Other unconventional platoon recognition indicators were proposed in recent years. For Lin (*68*) "the longest segment of consecutive vehicle groups with similar headways" constitutes a "platoon body." The standard deviation of intra-vehicle headway establishes a cutoff, i.e., whether a vehicle belongs to a platoon.

Some researchers sought to test their hypothesis that a factor of two to three can increase bottleneck capacity if vehicles are organized to cross the intersection in platoons with very short headways. In this regard, Lioris et al. (*69*) tried to predict the performance of a cooperative adaptive cruise control (CACC) enabled system, using three different delay estimation models and a simulation (hypothetical) case study, when saturation flows (on average) increase by any factor (up to 300%). Experimental results of platoons of self-organized short-headway CACC vehicles suggest that a saturation flow rate increase, indicates an increase in the demand by the same factor, with no increase in queuing delay or travel time, regardless of the control strategy applied. Simulation experiments were, however, limited to $V/C < 1$ conditions.

Recognizing a traffic controller's complex operational requirements, the adaptive control algorithm by Feng et al. (*70*) integrates various components, such as multimodal priority requests, platoon-based coordination requests, and normal vehicle-actuated control, when determining the sequence of signal phases and their durations. The reason for incorporating regular-vehicle actuation in the "connected" controller logic were the errors of the unequipped vehicle position

estimation algorithm, under low penetration rates. As the penetration increases, actuation may negatively affect the performance of the adaptive control algorithm since sufficient connected data are available to make better decisions.

To reduce computational burden and complexity on the controller, Jin et al. (*71*) proposed a reservation-based intersection management system with vehicle self-platooning. Each platoon's lead vehicle communicates with the intersection by sending the estimated earliest arrival and clearance time of its platoon. After the traffic controller confirms the reservation, the leader will design its trajectory and trajectories of its followers to meet the chosen criteria.

Connectivity-enabled platoon-based control strategies continue to emerge. Yang et al. (*72*) tested their switch or extend signal timing logic to reduce platoons' waiting time on a hypothetical four-phase isolated intersection, where segmentation of platoons was based on critical headway threshold. For a pre-determined phase sequence and before running each phase, waiting time are balanced against the time it takes to clear the last queued vehicle times the number of vehicles served if the phase is to stay active.

2.2.1. Joint optimization of vehicle trajectories and signal timing parameters.

Additionally, there have been limited efforts to design approaches that control signal operations and vehicle trajectories in an integrated manner. Sun et al. (*56*) designed a method which controls lane changing and car-following behavior while optimizing splits to maximize intersection capacity. Under the V2I framework and with full connectivity, Li et al. (*73*) developed and tested a joint vehicle trajectories and signal control parameters optimization algorithm. The approach was, timing plan enumeration based to account for certain restrictions, to reduce combinatorial

complexity even further and fundamentally was focused on trajectories optimization. In a simplified setup (two-phase, two one-way roads), the proposed method was tested under a variety of demand scenarios (undersaturated conditions) and showed modest improvements are achievable compared to vehicle-actuated control. Subsequently, the authors extended the concept to incorporate mixed traffic environments (connected, autonomous and regular vehicles) and real-time optimization of control parameters. Even in a basic intersection/signal/traffic configuration, the controller only decided whether to switch or extend the current phase (*74*). Similar work has been done by (*75*), where approximate dynamic programming found optimal traffic signal schedules, and optimal vehicle speed advice is given after the corresponding green has started, conditioned upon no queue had formed. Overall procedure was aimed at minimizing both delay and number of stops. Xu et al. introduced a cooperation method between signals and vehicles' speed (*76*). The proposed method optimizes actuated cycle lengths while vehicle speeds are optimized on a rolling horizon basis. The model minimizes the travel time. Vehicle control minimizes the consumption by optimizing the braking and engine power. However, the study only considers autonomous vehicles.

Some of the more recent studies propose frameworks for dynamic traffic management to optimize network level signal control decision variables and departure times of individual connected vehicles to determine their optimal routes. Signal control parameters are updated at every control interval, while departure times arise upon vehicle's request. The latter problem is solved as the shortest path problem, where links availability depends on the decisions previously taken by all the other vehicles (*77*). A similar solution was proposed for grid subnetworks by (*78*). The authors,

however, predict vehicle turning movement, i.e., travel direction according to a discrete probability distribution function and assume phase sequence and duration are fixed.

Control methods reviewed in this study, if tested, proved efficient only under light-demand scenarios, which is expected, considering the complexity of the problem and consequently computational effort required to solve the problem in a reasonable amount of time.

2.3. Analysis of network performance at path level in a connected environment

An elementary step in conducting a study is to identify a suitable measure or indicator of the state. Though detailed, knowing traffic conditions on every link of a city's network at every second is not practical. On the other end, network-level aggregated traffic conditions over an entire day will lack most of the details. There is a trade-off between the level of detail and the state reporting conciseness. However, an aggregate model approach that considers the traffic dynamics of a large urban area at the right aggregation level is promising for such studies. A Network Fundamental Diagram (NFD) or Macroscopic Fundamental Diagram (MFD) links space-mean flow, density, and speed of a large urban area. Many types of research and studies have indicated the existence of NFD (*1; 9; 20; 79-85*). One of the advantages of using NFD is its capability to capture congestion, breakdown, and recovery. When a facility reaches its capacity, flow and speed decrease with the addition of vehicles and traffic approaches gridlock. This phenomenon is also known as 'backward-bending' (*86; 87*) phenomenon or 'hypercongestion' (*88; 89*). NFDs describe the regional traffic state, mainly the gridlock phenomenon during peak hours. Hence, NFDs can be used to study more closely the dynamics of traffic flow. The concept of NFDs can also be

extended and applied to arterial roads. It has been recognized that the arterial fundamental diagram (AFD) is significantly affected by signal operations. Wu et al. (2011) explored the impacts of signal operations on the AFD. The authors concluded that the stable form of AFD is of great importance for traffic signal control because of its ability to identify traffic states on a signalized link (*90*). It has been established that NFD (AFD) should be treated as an essential indicator for traffic states under signal control or any other kinds of control. Given the potential of NFDs, this work will recognize them as the measure of the traffic state. An important aspect to consider is the selection of the right level of aggregation, i.e., the right level of network partition into smaller regions.

Clustering vehicle trajectories and discovering similar individual trips lead to a range of location-based service applications (*91; 92*). Applications of spatial clustering in transportation have received more attention in the last couple of years. Using individual trajectories, Palma et al. (*93*) proposed a speed-based spatiotemporal clustering method to identify frequently visited places. Chen et al. (*94*) developed a coherence expanding algorithm and adopted the absorbing Markov chain model to investigate the problem of discovering the most popular route through global positioning system (GPS) trajectories. Guo et al. (*95*) conducted a spatial clustering of massive GPS points to recognize potentially meaningful places and extract the flow measures of clusters to understand the spatial distribution and movements. Bahbouh et al. (*96; 97*) proposed a framework to identify on-demand corridors from origin-destination information. One of the aims of this work is to design a trajectory clustering algorithm to select a set of popular paths for further NFD exploration. Relevant literature shows that the spatial variability of vehicle density is a critical factor that affects the shape, the scatter, and the existence of a well-defined and stable MFD

(*22; 82*). Ji and Geroliminis (*98*) developed an algorithm to cluster heterogeneous traffic networks into homogeneous regions. Figure 2-1 illustrates the clustering results obtained based on this method. It can be observed that the resulting clusters are divisions of the network based on the geographical location of the links.

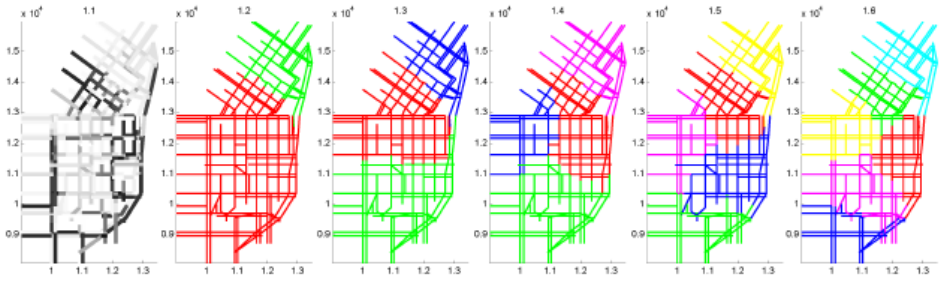


Figure 2-1 Cluster for heterogeneous transportation networks (source: Ji and Geroliminis 2011 (*98*))

This work will establish a clustering method based on the popular paths. Correspondingly, the clusters will not be obtained by factoring in the geographical positions and will be heterogeneous unlike the previous work (*98*).

The question that arises is *how information on the obtained clusters will be utilized*? Obtained clusters can be used to study the traffic flow pattern that can help an operator manage and control the network based on the strategies tailored to specific clusters' attributes. Management and control of multi-region NFD system can improve urban mobility, prevent overcrowding, and relieve congestion in cities. According to Geroliminis et al. (2012), the shape of NFD is not very sensitive to different demand patterns. This property is vital for control purposes because efficient active traffic management schemes can be developed without detailed knowledge of origin-destination (OD) tables.

Conversely, this effort is aimed at studying the paths and their Path Fundamental Diagrams (PFDs). Clusters revealed that a difference in the shapes does exist. Consequently, a custom-made control strategy corresponding to a cluster can be developed.

Another intersecting direction of research is determining the travel time distribution along the same lines is along the user-defined paths. There is a vast literature that considers link travel time as deterministic functions of traffic flows. Practitioners mainly apply this approach due to its simplicity and low computational burdens. However, link travel times are expected to be correlated over time and space and follow probability distributions that depend on that of the neighboring links. Numerous studies are conducted to estimate travel time joint distributions. Several stochastic network studies are performed to address the multivariate nature of the travel time distributions in stochastic pathfinding problems. Due to the analytical complexities, these problems mainly integrate the joint distributions with numerical integration or statistical sampling methods. In rest of the section, first, a review of the existing travel time distribution estimation studies is provided. Then, the existing efforts in integrating the travel time distributions along a set of links in path-finding problems are introduced.

This section contains a review of models and methods developed and practiced by researchers working on estimating travel time. Figure 2-2 presents a word cloud of the most frequent words in the reviewed documents. The literature review revolves around travel time distribution, correlation of travel time for links and paths (which consists of serial links), and models to capture correlation, different estimation and solution methods, random variable distributions, and data types. Based on the literature, a methodology is proposed to estimate travel time on paths using real-world trajectory data.

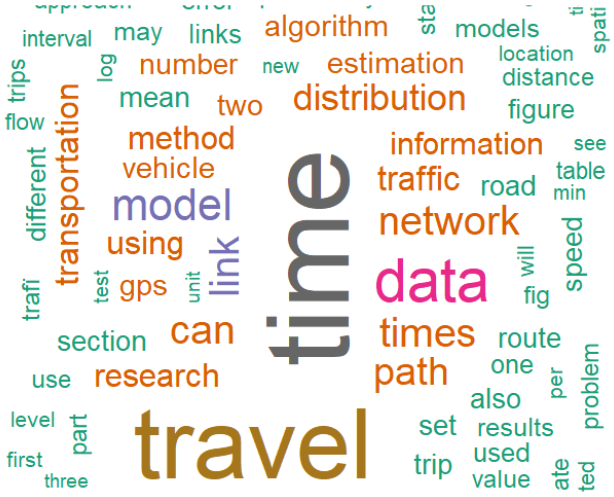


Figure 2-2 Frequent words in reviewed documents.

Models in the literature can be categorized as parametric and non-parametric. Further, these models either assume link travel times to be independent or dependent. Most of the researchers in the literature demonstrated the application of their methodology for a corridor for which they could obtain data. In the real world, one application of such studies is an online estimation, such as in Google Maps, Bing, etc., where a user makes an inquiry to estimate travel time over a path or a path and associated travel time between two points. However, in this project, the focus is on the offline estimation of travel times on user-defined paths. Extension of the developed methodology for the real-time application will be investigated in the future.

Depending on the needs of a particular application, travel time is determined either at the link level or the path level. A path is defined as a collection of serial links. The links typically have correlated travel times. However, some literature does not incorporate the correlation when determining path travel time modeled as a sum of link travel times. This results in a wrong estimate of the variance

of path travel time. As emphasized and established earlier, understanding and modeling a correct variance is an essential part of the study. “The spatial and temporal correlation of traffic data are intrinsic characteristics of road networks, which can be leveraged to solve both sensor data aggregation and optimal sensor placement problems in future smart cities.”(*99*) Accordingly, literature that assumes independence of link travel times is not discussed extensively below.

Studies concerned with the estimation of travel time distribution are mostly categorized into 1) link, 2) route, and 3) network levels (*100*). For these types of studies researchers have used data retrieved through different technologies, such as GPS (*101-107*), probe vehicles (*108*), computer simulation models (*109*), wireless signatures (*110*), dynamic Gaussian models (*108*), Bayesian models (*104; 105; 108*), map-matching (*111*), Markov chain, Gaussian mixture models (*103; 107*), gamma distribution, kernel density (*111*), and other parametric models (*101*) to solve the problem with expectation maximization (*103; 106; 107*) or likelihood maximization (*101; 104; 105*).

2.3.1. Link travel time distributions

Lu (*112*) develops a model to estimate corridor-level travel time distributions considering spatial and temporal correlations in urban freeway traffic data and probe vehicle data. This study compares different probability distribution shapes and found that lognormal distribution obtains the best fit.

Jenelius et al. (*101*) formulated a parametric model which takes into account the link attributes and operation conditions to estimate the link travel time rate. The model includes spatial moving average correlation structure for link travel times on the network and is demonstrated to be applicable at low market penetration rate (MPR) of GPS equipped vehicles.

A study by Hunter et al. (*108*) proposes a probe data-driven Gaussian model for a dynamic Bayesian network. The proposed method comprises three stages: identifying the stop and go behavior, computing the correlation in the behavior, and determining the correlations between link travel times. Based on the learning from the three stages, study infers the underlying travel time distribution.

2.3.2. Route travel time distributions

Several researchers have proposed methods to estimate route travel time distributions.

In (*106*) Ramezani and Geroliminis use probe vehicles travel times traveled along all links in an arterial route to estimate the route travel time distribution by a Markov chain bases method. Their method addresses the spatial correlations between successive links of a path by establishing a two-dimensional diagram for travel times of every two-consecutive links. These diagrams are then clustered to states with homogeneous travel times, and finally, route travel time distributions are estimated by multiplying the continuous Markov chain matrices assuming that the transitions between different links are conditionally independent.

A study by Isukapati et al. (*113*) constructs the corridor-level travel time distribution based on the segment-level travel time temporal and spatial distributions by adding together percentile-by-percentile values of the travel times. However, this technique may succeed only under specific conditions.

Eisele et al. (*114*) provide a method to estimate route travel time mean and variance as well as some link travel time distributional properties (e.g., confidence intervals).

Chen and Osorio (*115*) present an analytical method to approximate the standard deviation of path travel times and the results are compared with simulation models. However, the proposed method in this study is not validated with real-world data.

In (*116*), Chen et al. introduce a copula-based model to estimate the path travel time distribution on urban arterials. Estimated path travel time distributions are compared with those by the convolution and the empirical distribution fitting approaches. The presented results indicate the advantages of the copula method over those convolution and distribution-fitting methods. However, the proposed copula method requires a set of segment travel times as inputs that need to be estimated by another method to estimate link travel time.

Field data-based study by Kwong et al. (*110*) first identifies vehicles through their wireless signatures. For each vehicle, travel time was determined along the path as the difference between entry and exit time on the path. The study was conducted on a 1.5 km (0.93 miles) long signalized arterial road of San Pablo Avenue in Albany, California, under different traffic conditions. The study, however, does not provide any closed-form analytical solution to travel time nor does it provide a solution method. This study is merely basing results and presenting the obtained distribution from the collected data. However, Kwong et al. found that the travel time distribution is also a function of the lane. For example, the leftmost lane on a freeway, which is considered to be the fastest, is expected to have the lowest travel time or the rightmost lane can be the faster on an arterial road because no drivers are making left turns at intersections which do not have left-turn-only arrow signals. Furthermore, they conclude that the mean travel time is an insufficient indication of travel time reliability. This finding is very intuitive, but a data-based study gives further validation of the finding.

2.3.3. Network travel time distributions

Mahmassani et al. (*117*) use both simulated and real-world GPS trajectory data to model network travel time mean and standard deviation. Their results indicate that these two measures are highly positively correlated.

Hunter et al. (*108*) propose a method to estimate travel time distributions on a network using probe vehicle data. This study assumes that travel times on links are multivariate Gaussian distributed. However, assuming a single pattern for travel time distribution may lead to poor model fit. Reference (*101*) by Jenelius et al. develops a statistical framework to estimate the distribution of travel time on road network considering various factors including weather, speed limit, and seasonal effects. In this study, the travel time means, and variances are estimated as functions of these factors that expressed as explanatory variables. The link-level correlation structure is captured by a spatial moving average method using probe vehicle data.

Westgate (*104*) proposes two statistical methods to estimate travel time distribution on a road network using GPS-enabled ambulance data, and the model parameters are estimated by Markov chain Monte Carlo approaches. This study points out that GPS data readings are biased in the same direction and introduces a method to address this problem. Despite the helpful methodological insights into the network travel time distributions, the applied ambulances trip data may not completely represent regular urban traffic features.

Zheng et al. (*100*) present a network travel time distribution model based on Johnson curves, which can describe various travel time distributions patterns. The model parameters in this study are

estimated based on the percentiles in the field travel time data. Their results show that the applied method outperforms the widely used lognormal model by capturing the skewness in the field data.

Though the studies have developed and demonstrated promising models and solution techniques, it is essential to be aware of some pitfalls. Rahmani et al. (*111*) point out and eliminate the following potential sources of bias in the estimation of travel time distributions:

- Incomplete traversal of route:
    - It can underestimate the travel time of the path
- Influence of adjacent network;
    - If a vehicle bypasses a portion of the route, it can overestimate the travel time
- Non-uniform coverage of route:
    - Not all links will have the same number of observations
- Non-representative vehicle sample:
    - For example, taxi data does not represent regular vehicles; however, with some corrections, it can be used as a proxy for the traffic conditions

Autoregressive Integrated Moving Average (ARIMA) is another approach used in a set of studies that explicitly consider the correlation among the random variables. ARIMA is a popular technique that is used for prediction with correlated random variables. In the general traffic parameters that are predicted, for a short-term or long-term, are travel time (*118-122*), speed (*123-125*), and volume (*123; 126*). Given the importance of the correlation between spatial and temporal characteristics, studies on ARIMA are also reviewed, and a methodology around the technique is

formulated and presented later in the report. Use of ARIMA for prediction of traffic conditions is not a new concept (*123; 127-131*). These approaches invoke spatial and temporal relationship between random variables. As stated above, traffic state at a location gets affected by the neighbors and is correlated to the adjacent locations. Previous studies based on ARIMA involve the application of multivariate spatial-temporal autoregressive model (*123*), Space-Time Autoregressive Integrated Moving Average (STARIMA) model (*128*), spatial-temporal characteristics of flow on highways (*129*). Cheng et al. (*130*) compared complexities of models to compute correlations between links. Conventional methods to determine the spatial-temporal correlation include exploiting the information like neighboring links sharing a node (*127*), links that are approachable in a set number of time intervals (*123*), distance (*131*). Though studies develop new methods better than, regarding the computation efforts and the prediction accuracy, their past studies, assume stationary or quasi-dynamic spatial-temporal correlation.

Min and Wynter (*123*) developed a methodology using a multivariate model that incorporates the spatial-temporal effect from neighbors. They segment a day in different templates to represent different congestion conditions. For each template, a spatiotemporal matrix is determined, which is static after that. Hence, the methodology is not completely static; it is a quasi-dynamic methodology. Authors claim that their approach scaled well for the entire network and required fewer computation efforts. With the addition of a dynamic spatial-temporal correlation matrix, the methodology can be made applicable to real-time DTA models. Vlahogianni et al. (*132*) mention the challenge is to develop responsive algorithms and prediction schemes. It is essential to incorporate the inclement weather conditions or other non-recurring events in a predictive model. As much as it is needed, it is also difficult to see these events ahead of the time. For example, it is

challenging to predict an accident. With advances in the technology, we have better insight into the incoming weather condition. Incorporating the effect of non-recurrent conditions can provide more accurate predictions, enabling traffic management systems to maintain a sustainable level of service. Vlahogianni et al. suggest to include weather and incident responsive algorithms, enhancing the efficiency of online computations using artificial intelligence, and standardizing the requirements regarding the spatial and temporal data coverage in prediction methods. (*132*)

ARIMA models have been modified to incorporate effect from accidents or adverse weather to predict traffic states (*118; 119; 123; 126*). These studies do not explicitly mention incorporating the effects of weather in the reported results. It has been reported that the inclusion of rainfall (5 min data) in short-term travel time predictions could reduce forecasting inaccuracies and improve the model robustness. *(120-122; 124; 125)*

Studies revolving around travel time are not limited to transportation engineers. For example, study (*133*) was conducted by authors with computer science background. They identify the following major challenges in such studies:

- The sparsity of the data or low MPR of the GPS enabled vehicles
- Optimal aggregation level
- Real-time application of the estimation which needs an effective, scalable, and efficient method

Three-dimensional tensor is used to model travel time for different drivers on different links in different time slots. Using a dynamic programming solution an optimal value of the objective function is determined. The methodology is applied on 32,000 taxi trajectories collected over two

months. The authors introduced a tensor-based technique to stitch segments of a path together to obtain travel time. The method use vehicle trajectories to return travel time for a path inquired by a user. To obtain travel time along a segment where no or few trajectories are available, dynamic programming algorithm is used to fill the gaps where the objective is to find similar featured segments, learn the travel time pattern, and apply to the segment with missing values. This study is similar to an ARIMA technique presented in a tensor form.

CHAPTER 3. INTEGRATED SIMULATION FRAMEWORK DEVELOPMENT

To model and investigate the impact of connectivity on transportation systems as well as its impending applications, a framework to integrate micro aspects of individual vehicle interactions/behavior within a mesoscopic simulation tool was proposed. To be able to investigate facility or network-level effects of connectivity, vehicle communication and interactions are to be captured. On the other hand, to do so with adequate detail representation, microsimulation outputs are necessary, since it would be computationally challenging, for large-scale networks, to attempt to embed the level of detail necessary into a meso/macro environment. Therefore, microsimulation-based traffic stream variables are characterized and utilized to calibrate a mesoscopic simulation model. The meso-simulation model is then used to simulate and analyze facility/ network-level impact of connected vehicles.

Microscopic models are utilized to simulate the information exchange through the vehicle to vehicle (V2V) or vehicle to infrastructure (V2I) communication as well as individual vehicle (driving) behavior. Subsequently, a mesoscopic or macroscopic tool is needed that can scale and be computationally economical at a network level. This study integrates the microscopic aspect of V2V communication and driving behavior with the macroscopic aspect of dynamic traffic assignment at a network level.

3.1. Microscopic Model Calibration

To capture the effects of connectivity at network-level, it is critical to adequately model the collective effects of the interactions between connected and regular vehicles in the traffic stream. Accordingly, a microscopic simulation framework by Talebpour et al. (*134*) was adopted to identify the speed-density relationships at different MPRs of connected vehicles. This tool constructs different behavior modeling frameworks for regular and connected vehicles and can capture the collective effects of the interactions between them on traffic flow dynamics. The microsimulation outputs reflect this interdependent relationship and allow for mixed environment speed-density curves to be identified. These calibrated speed-density relationships are then used as the input to the mesoscopic simulation tool to simulate the network-wide effects of connectivity.

The details of these modeling frameworks and the calibration approach are discussed below.

3.1.1. Modeling Regular Vehicles

The acceleration behavior of regular vehicles is modeled based on the state-of-the-art car-following model of Talebpour et al. (*135*). This model, based on Kahneman and Tversky's prospect theory (*136*), recognizes that drivers exhibit different perceptions depending on the encountered traffic flow regime – i.e. congested versus uncongested.. Accordingly, two value functions were introduced, one for modeling driver behavior in congested regimes and one for modeling driver behavior in uncongested regimes. The uncongested (UC) traffic value function in this model has the following form:

$$U_{PT}^{UC}(a_n) = \frac{\left[w_m + (1 - w_m)\left(\tanh\left(\frac{a_n}{a_0} \right) + 1 \right) \right]}{2} \left[\frac{\left(\frac{a_n}{a_0} \right)}{\left(1 + \left(\frac{a_n}{a_0} \right) \right)^{\left(\frac{\gamma - 1}{2} \right)}} \right] \quad (3\text{-}1)$$

where U_{PT}^{UC} denotes the value function for the uncongested traffic conditions, $g > 0$ and w_m are parameters to be estimated, a_n is the acceleration chosen by the driver and $a_0 = 1m/s^2$ is used to normalize the acceleration. They proposed the following value function for the congested traffic condition:

$$U_{PT}^{C}(a_n) = \frac{\left[w'_m + (1 - w'_m)\left(\tanh\left(\frac{a_n}{a_0} \right) + 1 \right) \right]}{2} \left(\frac{a_n}{a_0} \right)^{\gamma'} \quad (3\text{-}2)$$

where U_{PT}^{C} denotes the value function for the congested traffic conditions, $g' > 0$ and w'_m are parameters to be estimated. At each evaluation stage, based on drivers' perception of their surrounding traffic condition, drivers employ the corresponding value functions to evaluate the gains from the chosen acceleration. Because the actual function applied by the driver is not known to the observer, the applicable regime is viewed as a latent state of the driver. Accordingly, the observer (analyst) can only formulate a probabilistic mechanism to predict the driver's latent behavior regime. This introduces a binary probabilistic regime selection mechanism into the evaluation stage, given by:

$$U_{PT}(a_n) = P(C).U_{PT}^{C} + P(UC).U_{PT}^{UC} \quad (3\text{-}3)$$

where U_{PT}, $P(C)$, and $P(UC)$ denote the expected value function, the probability of driving in a congested, and the probability of driving in uncongested traffic regime, respectively. The utility of each choice is calculated using the following equation:

$$Y = \beta K + \varepsilon \quad (3\text{-}4)$$

where K, β, and ε denote a vector of variables (see table 1), a vector of unknown parameters to be estimated, and an independent and identically Gumbel distributed error term, respectively. The Gumbel distribution for the error term results in a binary logit expression:

$$P(C) = \frac{e^{Y(C)}}{e^{Y(C)} + e^{Y(UC)}} = \frac{e^{[\beta(C) - \beta(UC)]K}}{1 + e^{[\beta(C) - \beta(UC)]K}} = \frac{e^{\beta' K}}{1 + e^{\beta' K}} \quad (3\text{-}5)$$

$$P(UC) = 1 - P(C) \quad (3\text{-}6)$$

Table 1 shows the calibration results for β¢. Note that it is assumed that drivers choose the acceleration value function that gives them the highest value for the observed acceleration. If the drivers choose a_n as their acceleration, they will gain utility value of $U_{PT}(a_n)$ if they do not get involved in a crash. If they are involved in a crash, their disutility corresponds to the crash seriousness term, $k(v, \Delta v)$, for the velocity of Δv an associated change in the velocity v. The

probability of crash involvement is calculated in the same way as proposed by Hamdar et al. (*137*). This work assumed crash involvement probability was normally distributed and calculated the probability of having negative spacing at the end of the evaluation period. The total utility function of acceleration is formulated as follow:

$$U(a_n) = (1 - p_{n,i})U_{PT}(a_n) - p_{n,i} w_c k(v, \Delta v) \quad (3\text{-}7)$$

where $p_{n,i}$ denotes the crash probability. Finally, to reflect the stochastic response adopted by the drivers, the logistic functional form specified by Hamdar et al. (*138*) is used to calculate the probability density function:

$$f(a_n) = \begin{cases} \dfrac{e^{\beta_{PT} U(a_n)}}{\int_{a_{\min}}^{a_{\max}} e^{\beta_{PT} U(a')} da'} & a_{\min} < a_n < a_{\max} \\ 0 & \textit{Otherwise} \end{cases} \quad (3\text{-}8)$$

where β_{PT} is a free parameter that reflects the sensitivity of choice to the utility $U(a_n)$. Note that this study adopted Talebpour et al.'s acceleration model to model car-following behavior in the absence of communication. Note that the parameters of this model are calibrated against Next Generation Simulation (NGSIM) data (*139*).

Table 3-1 Calibration results from microsimulation.

Data	Definition	Coefficients	Unites
K0	Model Constant	-37.8195	-
K1	Driver's speed	1.7535	m/s
K2	Average headway between driver i and her leaders in all lanes. A value of 9,999 is assigned if there is no leader.	0.0459	s
K3	Average Relative speed between driver i and her leaders in all lanes. A value of 999 is assigned if there is no leader.	0.3259	m/s
K4	Average Headway between driver i and her followers in all lanes. A value of 9,999 is assigned if there is no leader.	0. 0931	s
K5	Average Relative speed between driver i and her followers in all lanes. A value of 999 is assigned if there is no leader.	-1.0300	m/s
K6	Driver's average surrounding density. It is defined as the total density (over the number of lanes).	0.5911	Veh/km/lane

Acceleration model parameters used in this study are shown in Table 3-2 in with their typical values.

Table 3-2 Acceleration Model Parameters and Their Typical Values

Parameters	Typical Value
Sensitivity Exponents of the Generalized	$\gamma = 0.5$
Utility Asymmetry Factor for Negative Utilities	$w'_m = 2$
Weighing Factor for Accidents	$w_c = 40$
Logit Uncertainty Parameter (Intra-Driver Variability)	$\beta = 3$
Maximum Acceleration	$\alpha_{max} = 4\ m/s^2$
Minimum Acceleration	$\alpha_{min} = -8\ m/s^2$
Acceleration Range Considered Near Interaction Point	$a_0 = 1(m/s^2)$

3.1.2. Modeling Connected Vehicles

The acceleration behavior of connected vehicles is modeled based on the Intelligent Driver Model (IDM) (*140*). IDM specifies a following vehicle's acceleration as a continuous function of the vehicle's current speed (v_n), distance s_n to the leading vehicle, and the difference between the leading and the following vehicles' velocities (Δv_n). Perceptive parameters such as desired acceleration, desired gap size, and comfortable deceleration are considered in this model (*140; 141*):

$$a_{IDM}^{n}(s_n, v_n, \Delta v_n) = \bar{a}_n \left[1 - \left(\frac{v_n}{v_0^n} \right)^{\delta_n} - \left(\frac{s^*(v_n, \Delta v_n)}{s_n} \right)^2 \right] \quad (3\text{-}9)$$

$$s^*(v_n, \Delta v_n) = s_0^n + T_n v_n + \frac{v_n \Delta v_n}{2\sqrt{\bar{a}_n \bar{b}_n}} \quad (3\text{-}10)$$

where δ_n (Free acceleration exponent), T_n (Desired time gap), $\bar{a}_n$ (Maximum acceleration), $\bar{b}_n$ (Desired deceleration), s_0^n (Jam distance), and v_0^n (Desired speed) are parameters to be calibrated. s^* is the desired (safe) gap. Note that the braking term in the IDM is designed to preclude crashes in the simulation.

3.2. Mesoscopic Model Calibration

Micro-simulation setup accounts for the difference in behavior as well as the interactions between the two types of vehicles. This research identifies a mixed traffic environment speed-density relationship based on the microscopic simulation results. To translate traffic flow dynamics of a connected environment from the microscale to the mesoscale, this micro speed-density relationship is applied on a mesolevel. Correspondingly, a 5.5 miles long highway segment on Interstate 290 in Chicago is modelled with the above-mentioned microscopic simulation tool at different MPRs of connected vehicles and different speed limits. Figure 3-1 shows the geometric characteristics of this highway segment.

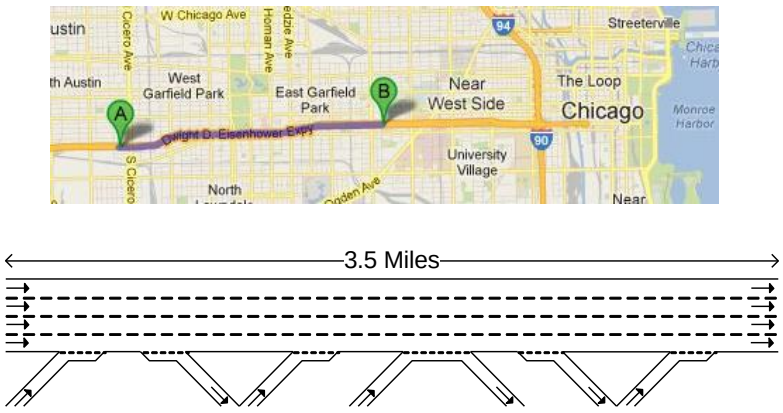


Figure 3-1 Geometric characteristics of the selected segment in Chicago, IL.

Figure 3-2 shows the schematic of the calibration approach based on the microscopic simulation results. Speed-density curves are identified based on multiple simulation runs (with different initial random seeds) until convergence is achieved. Once the model is calibrated, it will form the basis for the mesoscopic simulations.

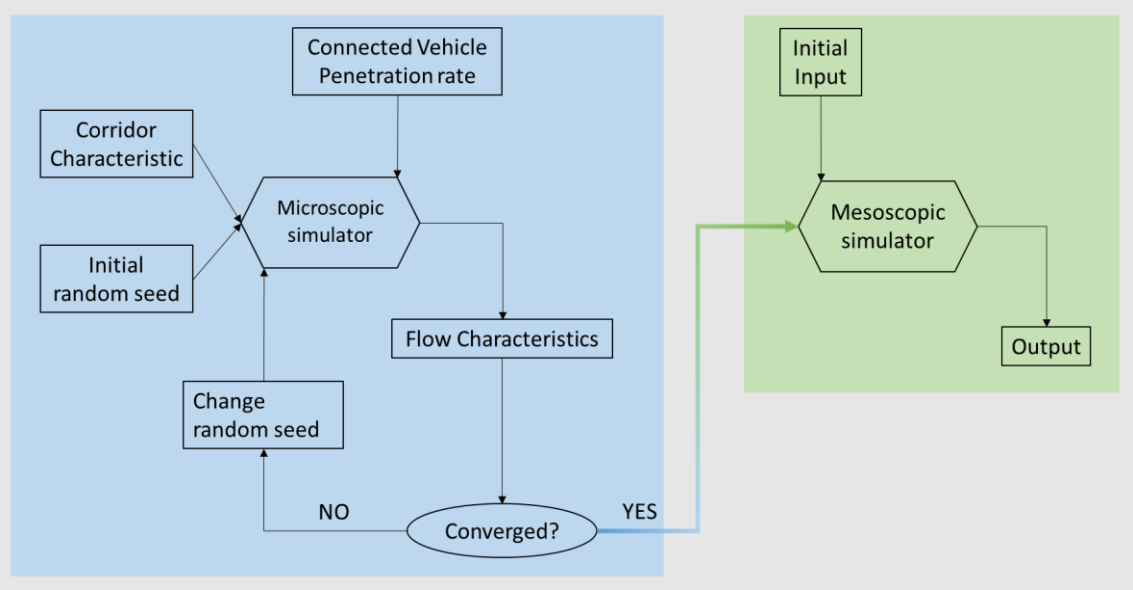


Figure 3-2 Schematic of the Calibration and Simulation Frameworks.

Four different MPRs of connected vehicles are selected (i.e., 0%, 10%, 50%, and 90%) along with four different speed limit values (i.e., 15, 35, 45, and 55 mph). Figure 3-3 shows the calibrated speed density curves. At lower speed limits, higher MPRs delay flow breakdown and the speed drop is observed at higher densities. At higher speed limits, however, the breakdown occurs at similar densities overall MPRs. Moreover, regardless of the speed limit, as the MPR of connected vehicles increases, higher speeds can be achieved for the same density value. This is particularly noteworthy in the congested regime of the speed-density curves.

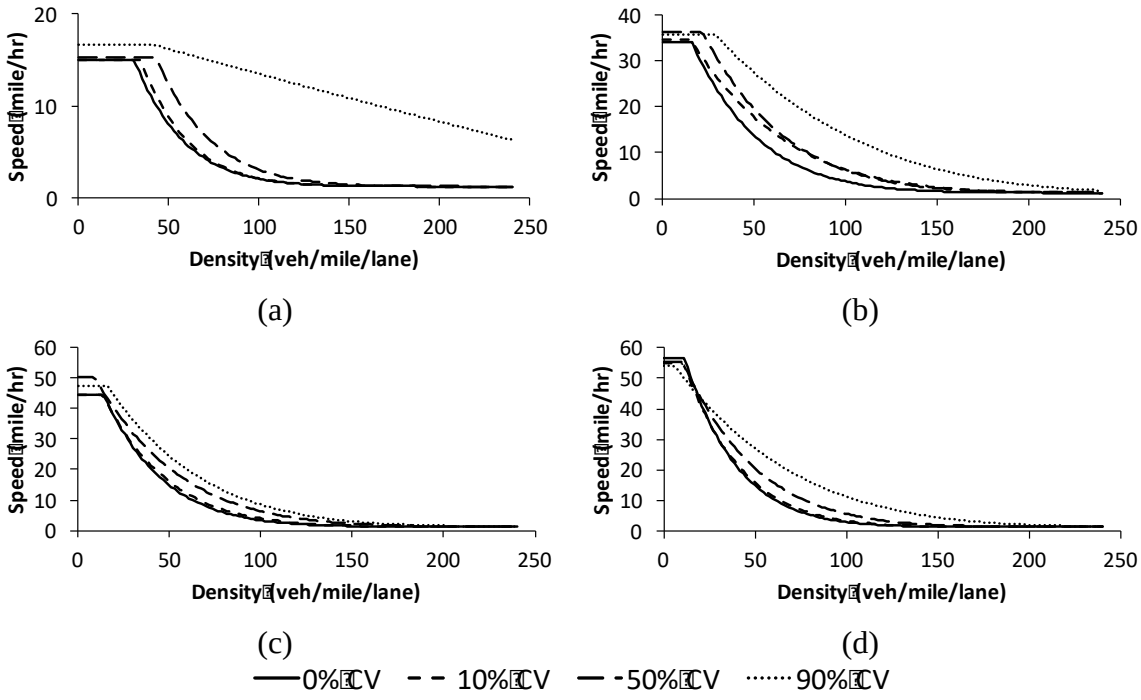


Figure 3-3 Speed-Density Curves at Different Market Penetration Rates of Connected Vehicles for Different Speed Limits: (a) 15mph, (b) 35mph, (c) 45mph, and (d) 55mph.

3.3. Measuring NFD and travel time variability

In practice, link-based measurements are computed to obtain the desired network-wide relationships among flow, density, and speed. Connected vehicle technology has the potential to change this widespread practice by providing information about individual vehicles throughout the network. Such information can be translated into vehicle trajectories and can be used to enhance the accuracy of measuring network-wide relationships as shown by Saberi et al. (*1*). They showed that trajectory-based measurements of network-wide traffic flow variables is indeed possible and provides a more accurate estimation of network-wide flow, density, and speed. In this study, trajectory-based measurements are adopted to obtain various variables and performance measures.

Following the approach proposed by Mahmassani et al. (*83*), NFDs are obtained based on trajectory-based calculations as follows:

$$Q(\omega) = \frac{d(\omega)}{L_{xy}(\omega).\Delta t} \quad (3\text{-}11)$$

$$K(\omega) = \frac{t(\omega)}{L_{xy}(\omega).\Delta t} \quad (3\text{-}12)$$

$$V(\omega) = \frac{d(\omega)}{t(\omega)} \quad (3\text{-}13)$$

Where

$Q(\omega)$: Average Network Flow

$K(\omega)$: Average Network Density

$V(\omega)$: Average Network Speed

$d(\omega)$: Total distance traveled by all vehicles in shape ω

$t(\omega)$: Total time traveled by all vehicles in shape ω

$d(\omega)$: Total distance traveled by all vehicles in shape ω

$L_{xy}(\omega)$: Total length (in lane mile) of the network on x-y plane associated with the shape ω

Δt: Time-height of the shape ω

Moreover, following the approach by Mahmassani et al. (*83*), in order to study the travel time variability, distance-weighted mean (μ) and standard deviation (σ) of the individual travel time rates (i.e., travel time per unit distance) are computed as follows:

$$m = \frac{\sum_{i=1}^{n} d_i t_i'}{\sum_{i=1}^{n} d_i} = \frac{\sum_{i=1}^{n} t_i}{\sum_{i=1}^{n} d_i} = \frac{1}{\bar{u}_s} \qquad (3\text{-}14)$$

$$s = \sqrt{\frac{\sum_{i=1}^{n} d_i (t_i' - m)^2}{\sum_{i=1}^{n} d_i}} \qquad (3\text{-}15)$$

Where

i: Vehicle Index

n: Number of Vehicles

d_i: Travel Distance of Vehicle i

t_i': Travel Time Rate of Vehicle i (Min/Mile)

t_i: Travel Time of Vehicle i

These concepts and equations make the basis for performance measures throughout the dissertation.

CHAPTER 4. IMPACT OF CONNECTED VEHICLES

There is a growing need to improve the mobility and safety of transportation systems and cities are undergoing constant changes to accommodate this need. However, cities face limited resource availability when expanding the existing road network. To address the challenge, cities are incorporating innovative technologies into their current operational practices. Connected Vehicles technology, as one of the latest technologies in surface transportation, provides the opportunity to create a connected network of vehicles and infrastructure. In this network, individual vehicles can communicate with each other through Vehicle-to-Vehicle (V2V) communications. Moreover, individual vehicles are connected to infrastructure and the Traffic Management Center (TMC) through Vehicle-to-Infrastructure (V2I) communications. The real-time information provided by this technology can improve drivers' operational, tactical, and strategic decisions; thus, improve mobility, safety and reduce emissions and energy consumption of existing transportation systems.

4.1. Numerical Experiments Results

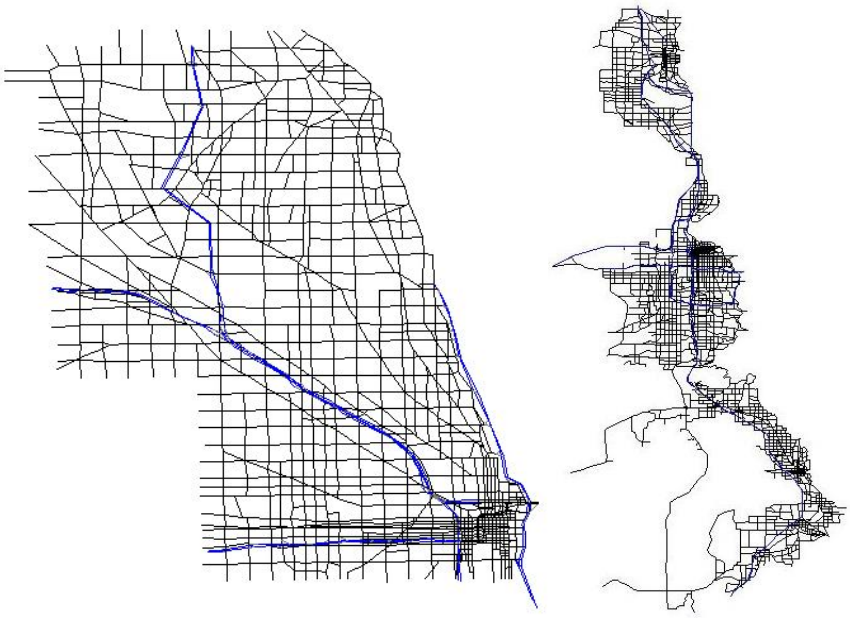
DYNASMART (**DY**namic **N**etwork **A**ssignment-**S**imulation **M**odel for **A**dvanced **R**oad **T**elematics), a dynamic traffic assignment simulation tool, was used to simulate traffic dynamics at the network level. The tool includes a dynamic network analysis and evaluation. One of the advantages of the tool is its capability to model individual vehicles' movements with different driver behavioral characteristics under various information guidance systems (*142*). Further details on the capabilities of the tool can be found elsewhere (*143*). The tool is being used in several

region-wide traffic management projects and can be made available to the public through the McTrans Center.

In this chapter, results from numerical experiments for Chicago and Salt Lake City modeled networks are presented. Figure 4-1 and Table 4-1 present the characteristics of these networks. The Chicago testbed network was extracted from the entire Chicago Metropolitan Area network to enhance the estimation and prediction accuracy during the implementation procedure. The testbed includes Chicago downtown area located in the central part of the network, Kennedy Expressway of I-90, Edens Expressway of I-94, Dwight D. Eisenhower Expressway of I-290, and Lakeshore Drive. The testbed is bounded on the east by Michigan Lake and on the west by Cicero Avenue and Harlem Avenue. Roosevelt Road and Lake Avenue bound the testbed from south and north, respectively. A full-size Salt Lake City network, consisting of Ogden-Salt Lake City-Provo area, was also used in this study. Details related to network configuration and calibration can be found elsewhere (*144; 145*). These two networks will represent testbeds throughout this research effort.

Three demand levels (low, medium and high) are simulated at four different MPRs of connected vehicles (i.e., 0, 10, 50 and 90%); the remainder of vehicles are conventional vehicles with no connectivity. The typical demand level is calibrated in accordance with the historical static O-D matrix and time-dependent traffic counts on observational links. Different demand levels refer to the demand pattern that has the same distribution as the average level but different fractions of vehicles. Low demand level has 40%, medium demand level has 70%, and the high demand level has 90% of the average level vehicles. The entire planning horizon or simulation period was divided into a 5-min interval series. At each 5-min interval, generalized network-wide traffic flow

variables, based on the extended Edie's definition, were computed using Equations (3-11) to (3-15). To report travel time reliability, mean and standard deviation of travel times were calculated using Equations 12 and 13.



(a) (b)

Figure 4-1 Schematic Diagram of (a) Chicago and (b) Salt Lake City.

Table 4-1 Network Configurations

Network	Chicago	Salt Lake City
Number of nodes	1,578	8,022
Number of links	4,805	17,947
Number of vehicles	805,275	937,483

Demand duration (hours)	24	15

Results for the Chicago area are presented in Figure 4-2 through Figure 4-5. Figure 4-2a presents the fundamental diagrams for different connectivity levels in Chicago's highway network at a low demand level. In this figure, an increase in connectivity results in a decrease in density and increase in flow. At the same density level, connectivity enables vehicles to move at a higher flow rate. The breakdown density and flow are increased, which results in a higher throughput in the highway network. Additionally, connectivity can reduce the maximum density observed in the network. For MPRs of up to 50%, the flow rate increases with connectivity, while the maximum density experienced in the network remains the same. At a 90% connectivity level, in addition to the flow rate increase, maximum density is reduced, as well. This indicates that connectivity can increase network's capacity and throughput at low demand levels.

Figure 4-3a shows the fundamental diagrams for different connectivity levels in Chicago's highway network at a medium demand level. Compared to Figure 4-2a, the impact of connectivity increases. The flow rate increases with the increase in connectivity, for the same density levels. Moreover, at 50% or higher MPRs, higher flow rates are achieved at the same density, which can result in faster recovery after breakdown. Note that at high MPRs, a clear hysteresis loop is observed at both low and medium demand levels (see Figure 4-2a and Figure 4-3a).

Figure 4-4a shows the fundamental diagrams for different connectivity levels in Chicago's highway network at a high demand level. In this case, not all vehicles reach their destinations due to gridlock, making corridors highly dense. In fact, high demand operational conditions reveal the

effects of connectivity on network-wide traffic flow dynamics under extreme conditions. Significant breakdown and scatter in the NFDs are observed at all connectivity levels. In extreme conditions, low MPRs of connected vehicles do not result in the improvement in flow and/or density, and the flow-density relationship remains almost unaffected (MPR of 10%). At a connectivity level of 50%, the network starts to experience an increase in the flow rate for the same density level, as compared to 0% and 10% MPR cases. When the MPR is 90%, highway network performs at its best, as the highest flow rate is achieved. At this level, highways reveal a small hysteresis loop and finally break down to never recover again, still higher throughput is achieved as compared to the other lower MPRs. Therefore, connectivity enables the network to enhance its throughput rate, even in extreme conditions.

Figure 4-2 through Figure 4-4 present the effects of connectivity on the network-wide flow-density relationship in Chicago's full network at low, medium, and high demand levels. At a low demand level (Figure 4-2b), as the MPR increases, the network experiences a greater reduction in maximum density. At the same time, an increase in flow rate for the same density level is observed. Connectivity is increasing the throughput while controlling the density in the network. At a medium demand level (Figure 4-3b), closed loop hysteresis is observed at all connectivity levels. The width of the hysteresis loop reduces with the increase in connectivity. This implies that connectivity facilitates the recovery. At a high demand level (Figure 4-4b), grid-lock is observed, making the corridors highly dense. Network-level operational characteristics exhibit similar flow-density relationships as shown in Figure 4-2 and Figure 4-3. It is the extent of the relationship that gets affected. Although, the extent of the relationship is altered. With the increase in connectivity,

the maximum density decreases. Moreover, connectivity increases the throughput of the entire network.

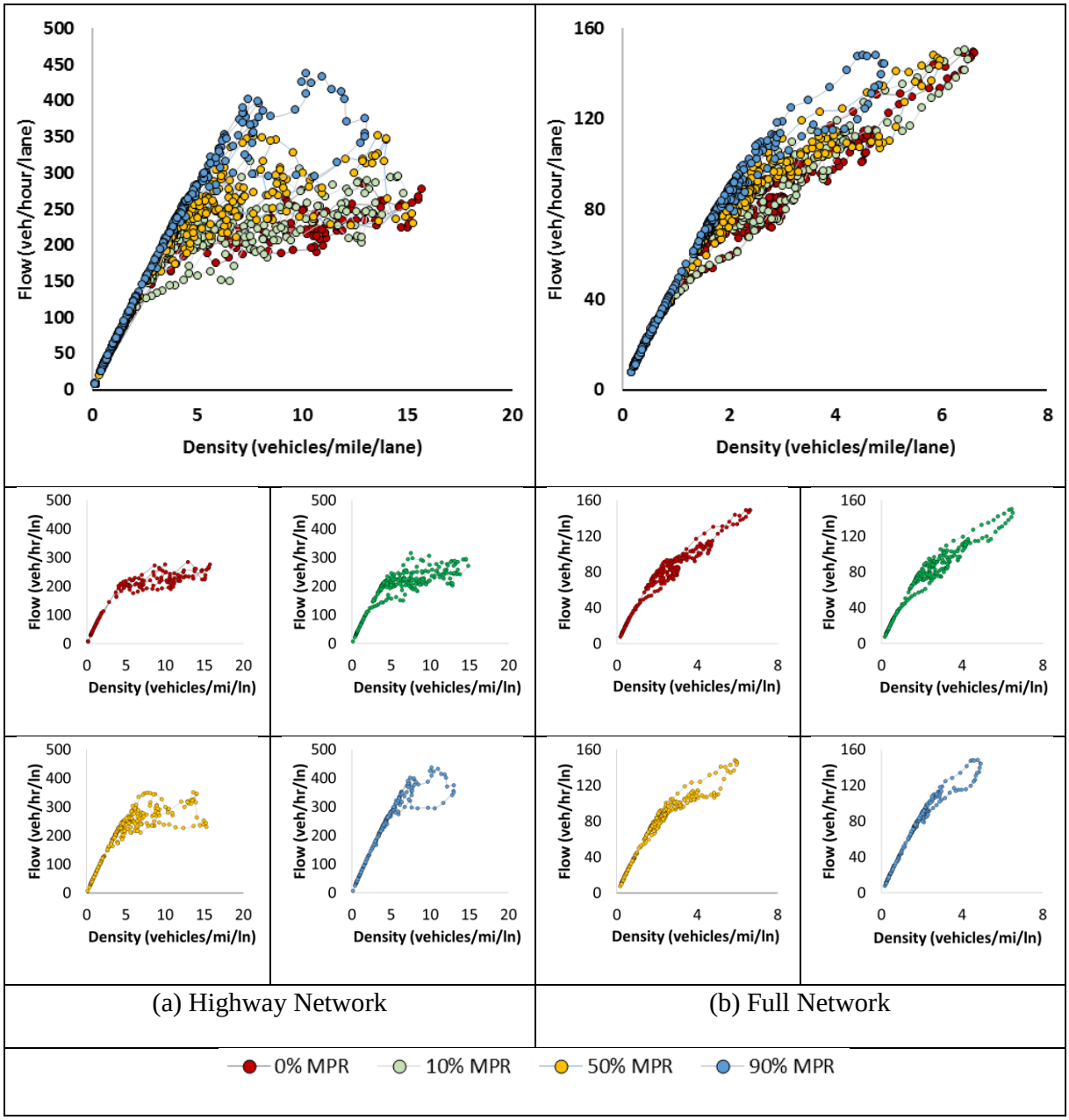


Figure 4-2 Fundamental Diagram for Chicago (a) highway network, and (b) Full network at different market penetration rates of connected vehicles under low demand.

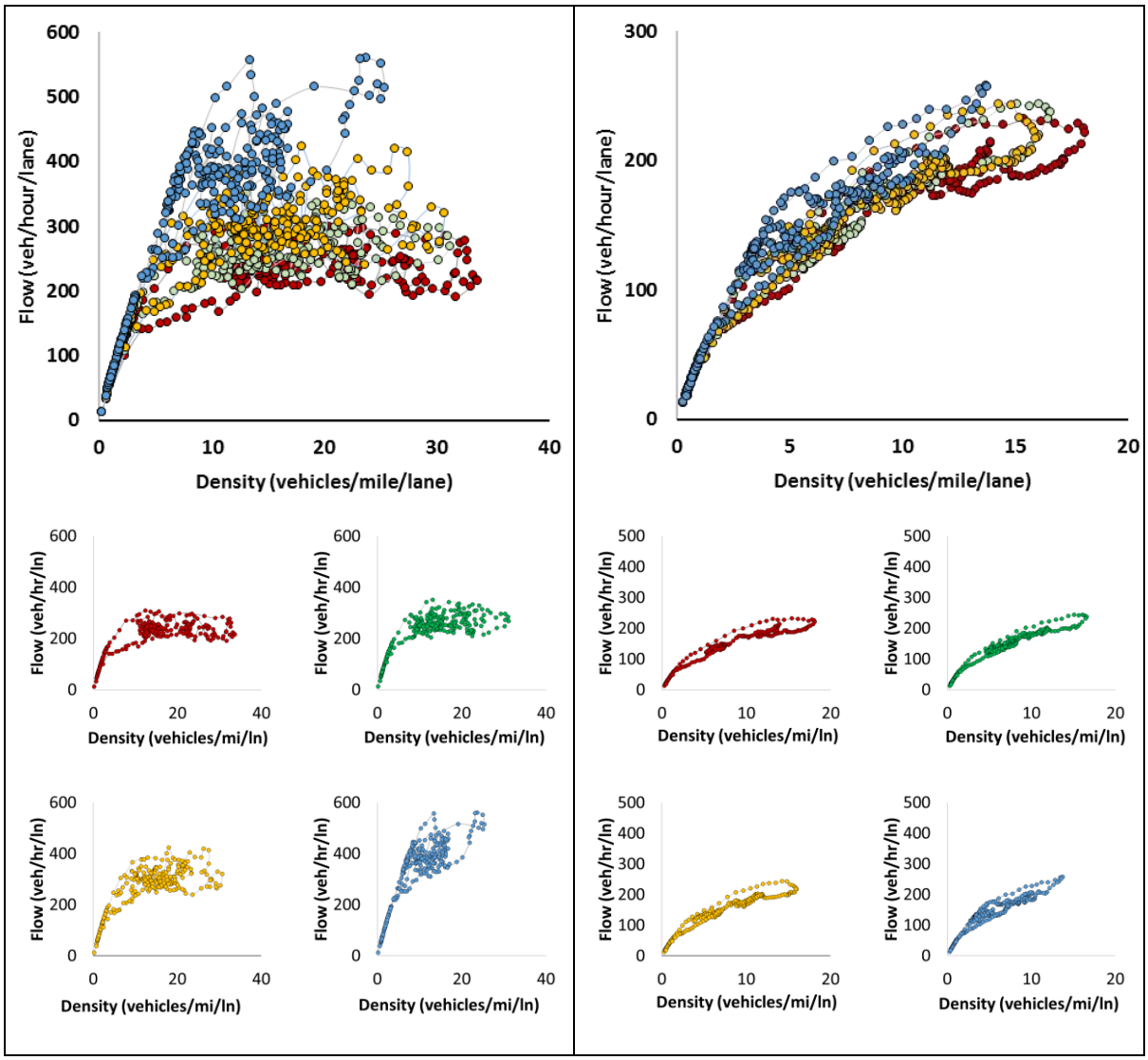
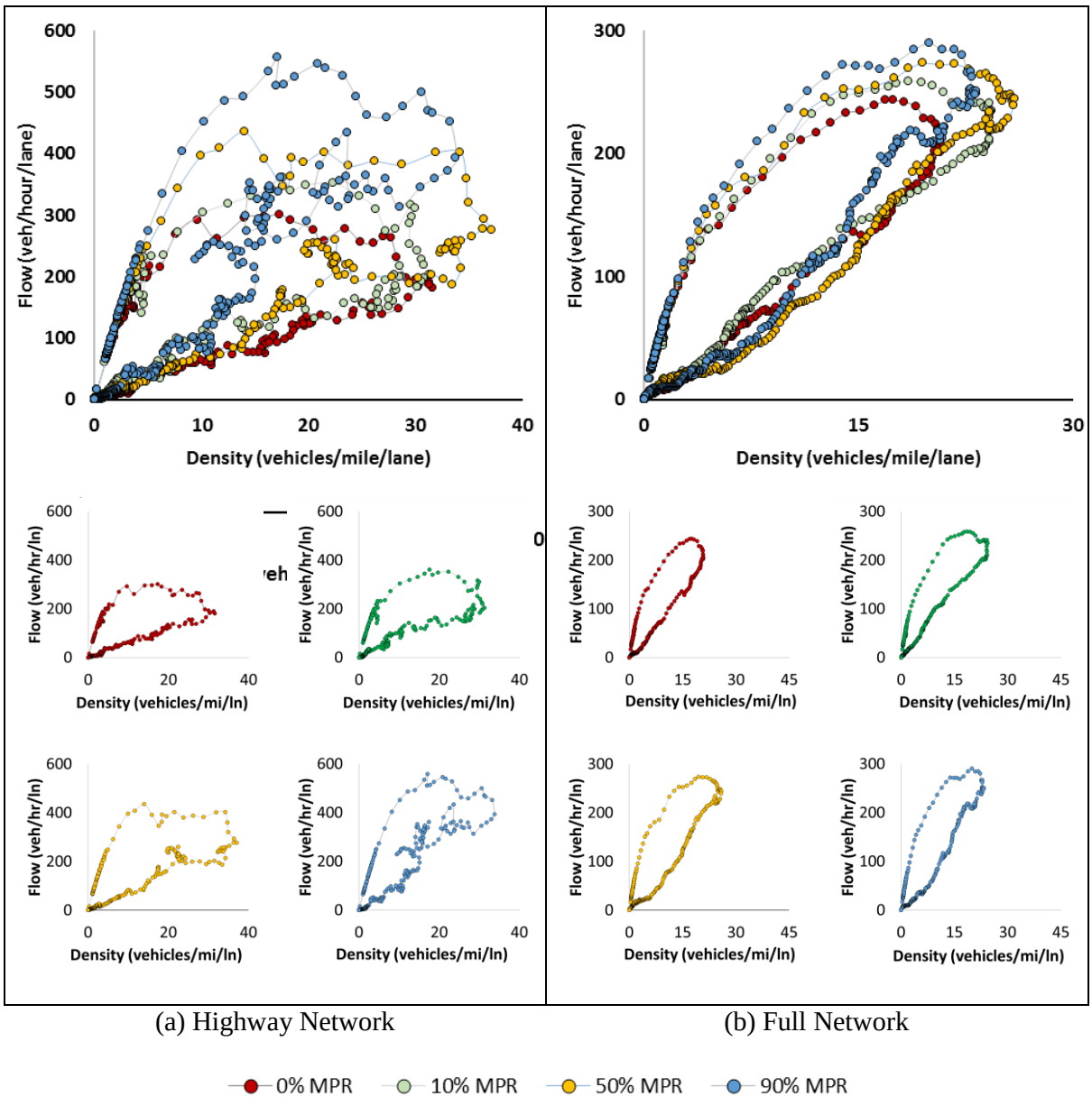


Figure 4-3 Fundamental Diagram for Chicago (a) highway network, and (b) Full network at different market penetration rates of connected vehicles under medium demand.

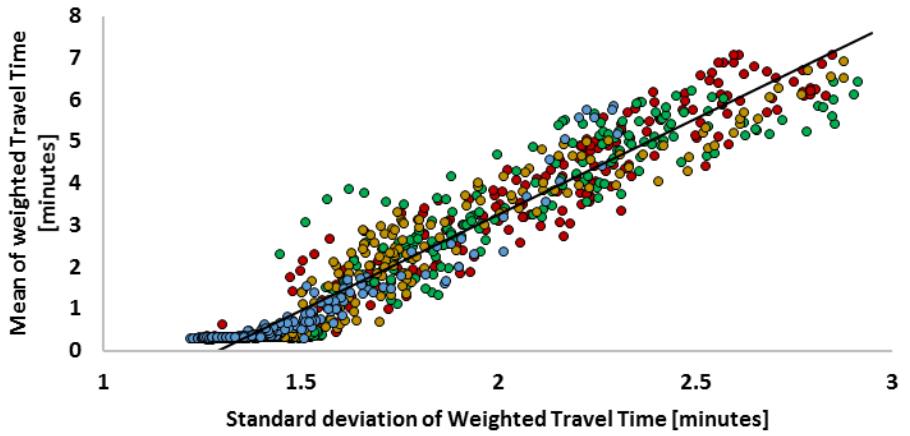


(a) Highway Network (b) Full Network

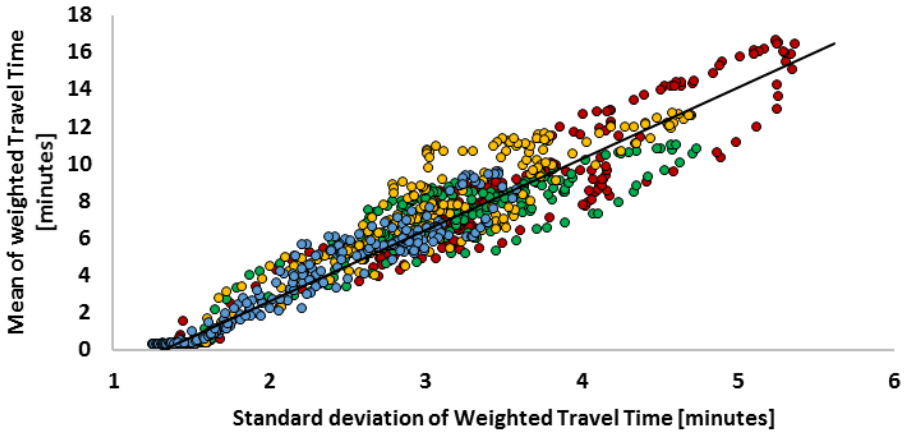
Figure 4-4 Fundamental Diagram for Chicago (a) highway network, and (b) Full network at different market penetration rates of connected vehicles under high demand.

The results presented in Figure 4-5 show the effects of connectivity on the travel time reliability at low, medium, and high demand levels. For all connectivity levels, weighted mean and standard deviation of travel time follows a similar linear trend but the extent is different. This finding

indicates that the linear relationship that Mahmassani et al. (*83*) found between the mean and standard deviation of travel time at the network-level is still applicable to the connected driving environment. At a low demand level (Figure 4-5a), with an increase in connectivity, the size of the distribution gets smaller, which implies that travel time variability decreases. Additionally, there is a slight reduction in the standard deviation for the same travel time with an increase in connectivity. This implies that travel time reliability increases as MPR of connected vehicles increases. At the medium demand level (Figure 4-5b), the impact of connectivity is more prominent than at the low demand level. Similar to the low demand case, with an increase in connectivity, the extent of the distribution reduced, but this reduction is more significant than in the low demand case. At the high demand level (Figure 4-5c), the weighted mean and standard deviation of travel time, follow the same linear pattern with different extents. With an increase in the MPR of CVs, the maximum mean and standard deviation of travel time decreases, and the standard deviation decreases for the same mean travel time. At all demand levels, the network is performing better and more reliably with the increase in the MPR of CVs, while following the same linear travel time variability relation.


Mean of weighted Travel Time [minutes]
0
1
2
3
4
5
6
7
8
1
1.5
2
2.5
3
Standard deviation of Weighted Travel Time [minutes]

(a)


Mean of weighted Travel Time [minutes]
0
2
4
6
8
10
12
14
16
18
1
2
3
4
5
6
Standard deviation of Weighted Travel Time [minutes]

(b)

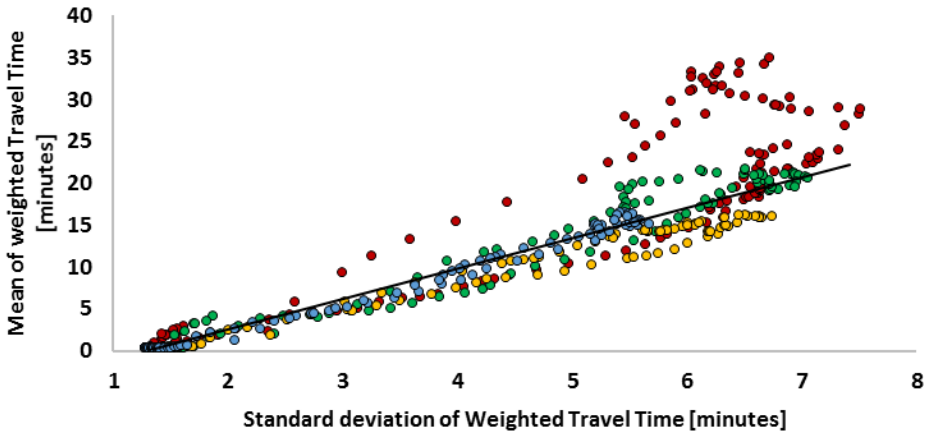

Mean of weighted Travel Time [minutes]
0
5
10
15
20
25
30
35
40
1
2
3
4
5
6
7
8
Standard deviation of Weighted Travel Time [minutes]

Figure 4-5 Travel time variability measures for Chicago with (a) low demand. (b) medium demand, and (c) high demand.

Results for the Salt Lake City area are presented in Figure 4-6 through Figure 4-9. Overall, the results show similar trends as in the Chicago network, but at a different scale. Figure 4-6a shows the fundamental diagrams of the Salt Lake City's highway network at the low demand level, different connectivity levels. In this figure, the highway network experiences a congested regime, which is prevented with advent of CVs. When the MPR is 90%, the breakdown is eliminated. Figure 4-7a shows the network-wide flow-density diagram at the medium demand level. In this figure, the highway network is highly congested, which ultimately results in a breakdown. However, high congestion levels are not observed at 90% MPR of CVs. Hence, as the MPR of connected vehicles increases, throughput increases as well. Figure 7-8a presents the network-wide flow-density diagram at the high demand level. Results are similar to those at the medium demand level, except the network does not reach jam density and shows recovery only at a 90% connectivity. However, at 50% MPR, the network-wide flow-density relationship is similar to the one observed when at 10% MPR of CVs. Throughput, however, still increases as the MPR of CVs increases. Hence, connectivity can help improve system's performance by increasing throughput rate and lowering the maximum density attained.

Part b of Figure 4-6 through Figure 4-8 presents the effects of connectivity on network-wide flow-density relationships for the full network at low, medium, and high demand levels in Salt Lake City. At a low demand level (Figure 4-6), as the MPR increases, the network experiences the reduction in maximum density and the increase in flow rate at the same density level. With only regular vehicles in the system, the whole network experiences an unrecovered breakdown, whereas

with connectivity, the system realizes improved flow conditions. At a medium demand level (Figure 4-7b), the system experiences considerable congestion even at low MPRs (the network experiences a density level of around 180 vehicles/mile/lane, which is an indication of gridlock). At high MPRs, however, the network recovers from the breakdown. When the MPR is 50% and 90%, the network flow increases and at in case of a 90% MPR, the system forms a closed hysteresis loop. At a high demand level (Figure 4-8b), the network only recovers when 90% of vehicles are connected, and the flow-density relationship is similar to the ones observed at 0%, 10% or 50% connectivity levels.

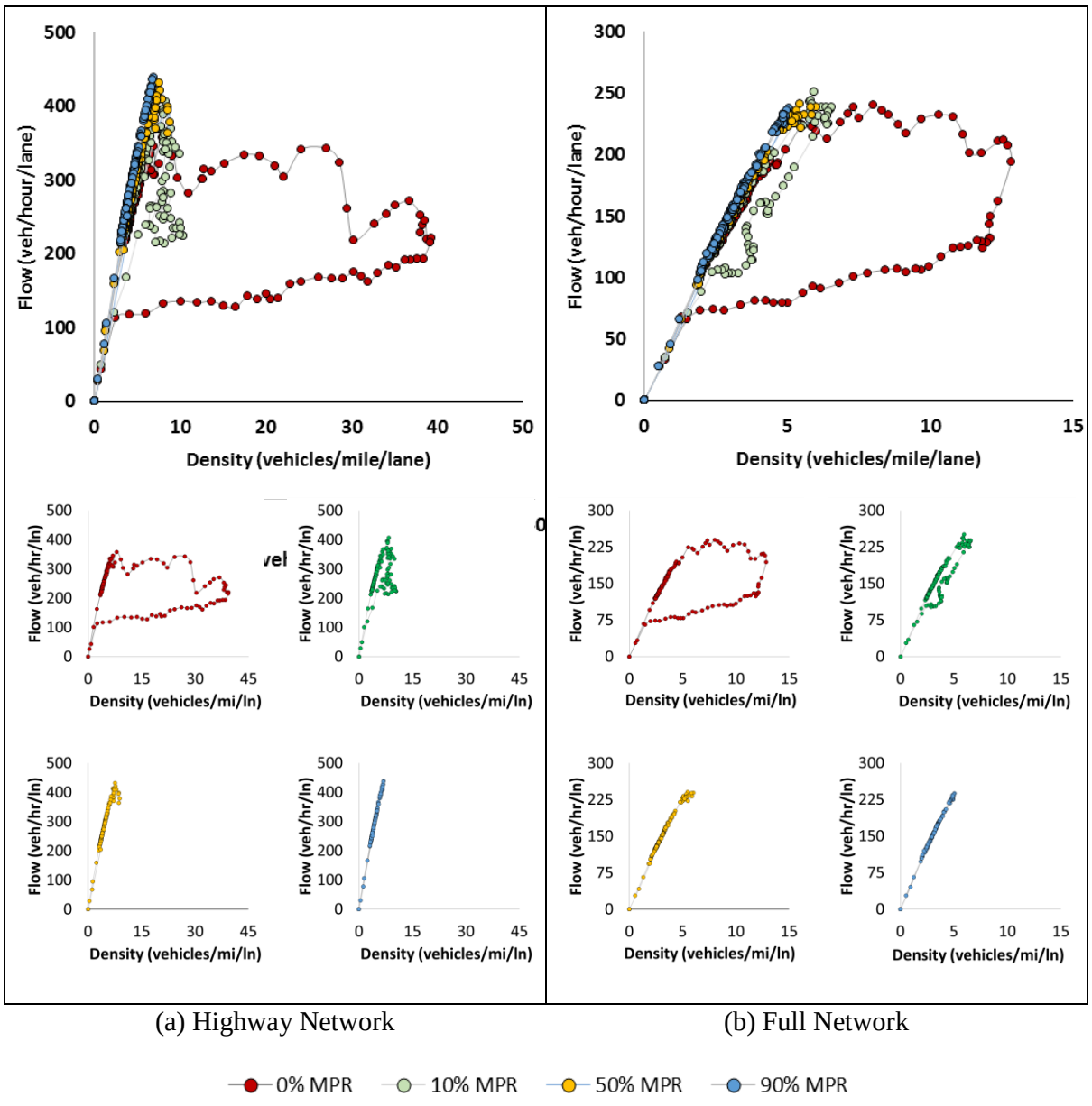
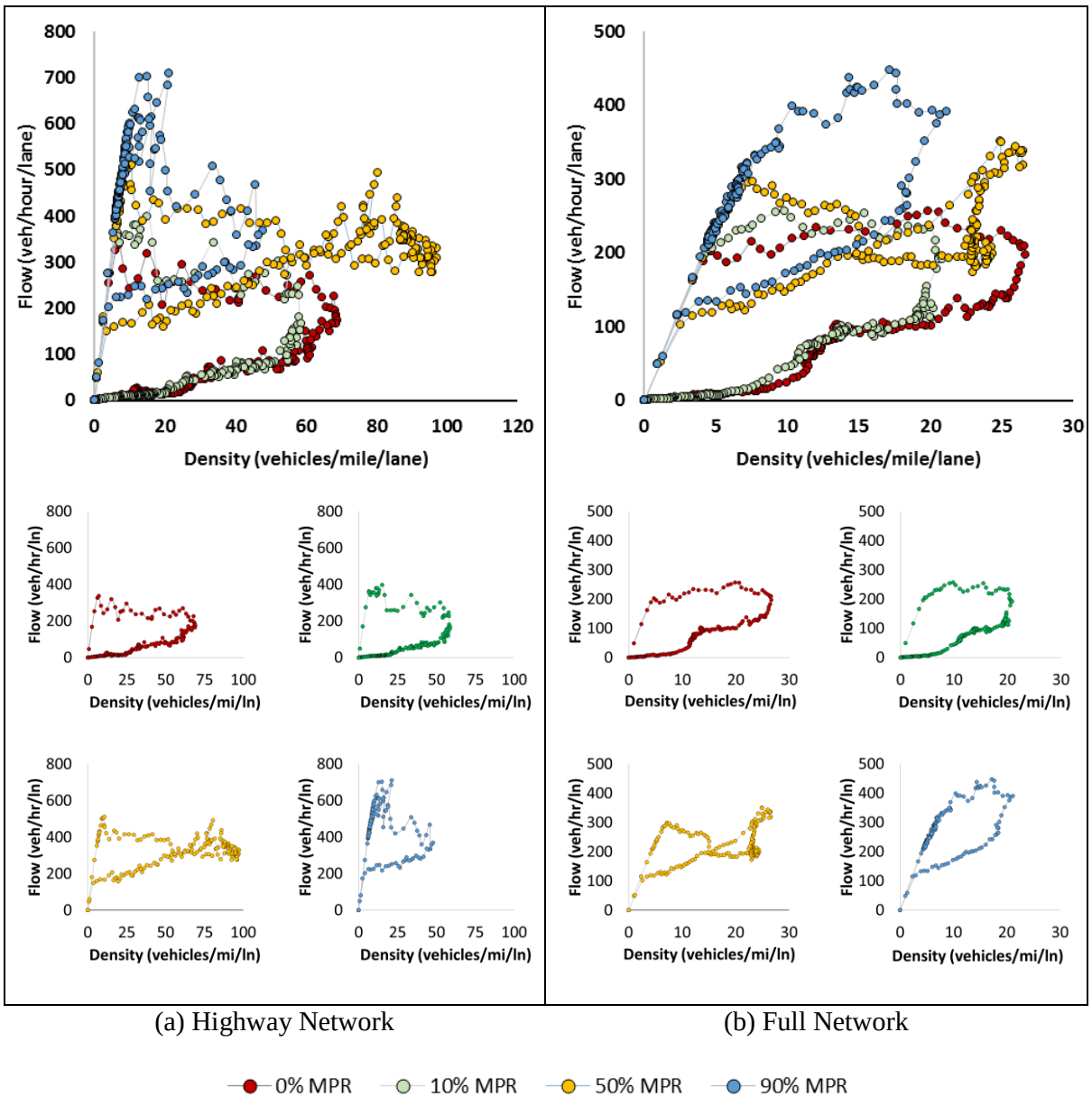


Figure 4-6 Fundamental Diagram for Salt Lake City (a) highway network, and (b) Full network at different market penetration rates of connected vehicles under low demand.



(a) Highway Network (b) Full Network

Figure 4-7 Fundamental Diagram for Salt Lake City (a) highway network, and (b) Full network at different market penetration rates of connected vehicles under medium demand.

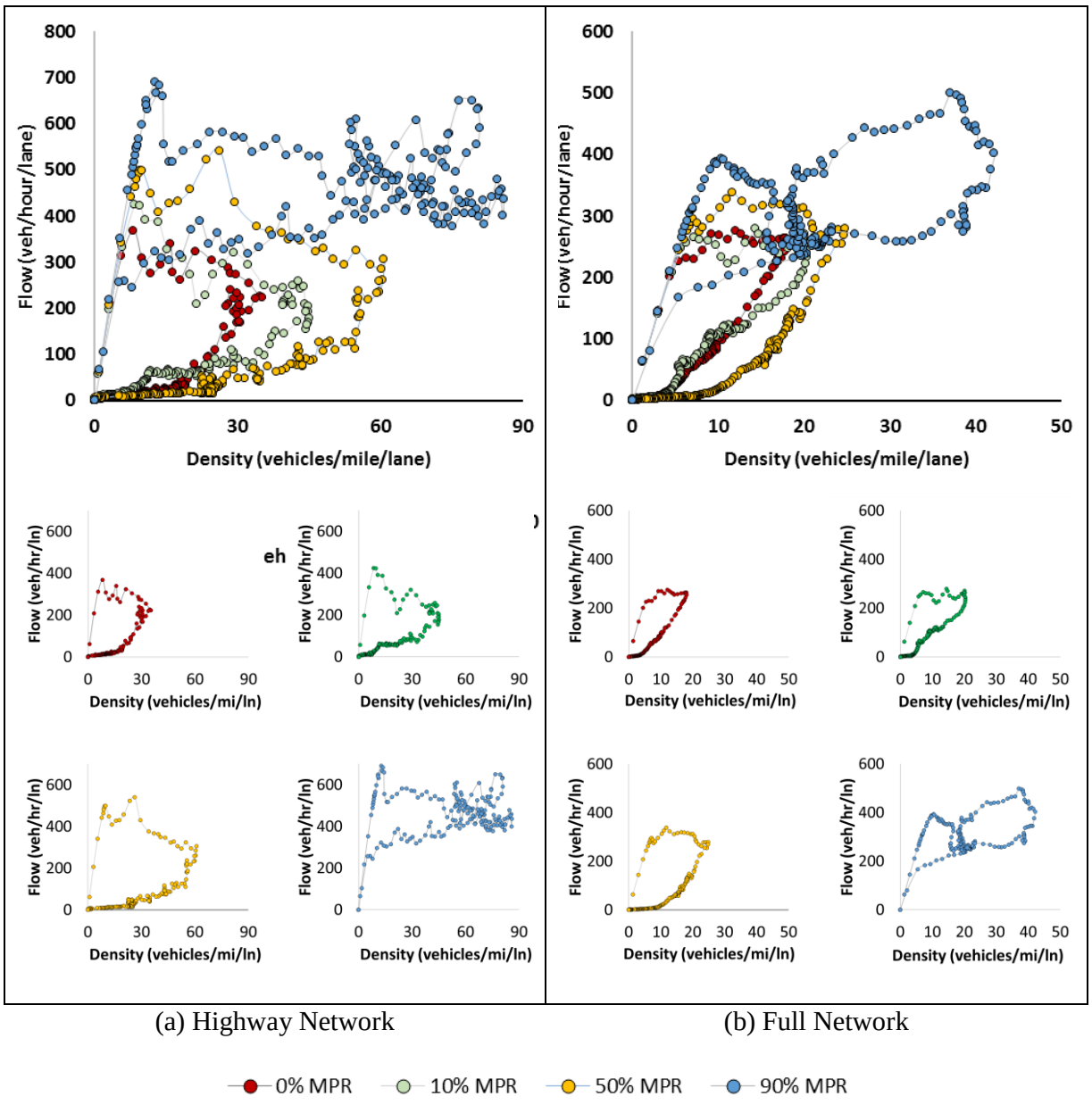
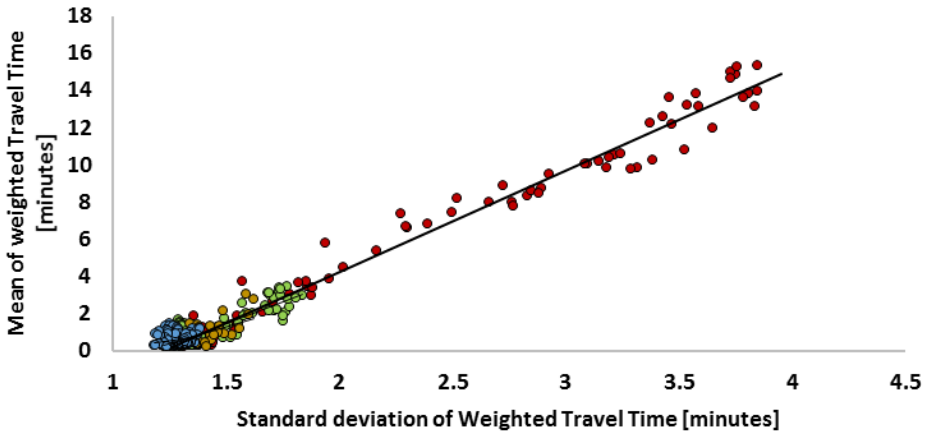


Figure 4-8 Fundamental Diagram for Salt Lake City (a) highway network, and (b) Full network at different market penetration rates of connected vehicles under high demand.

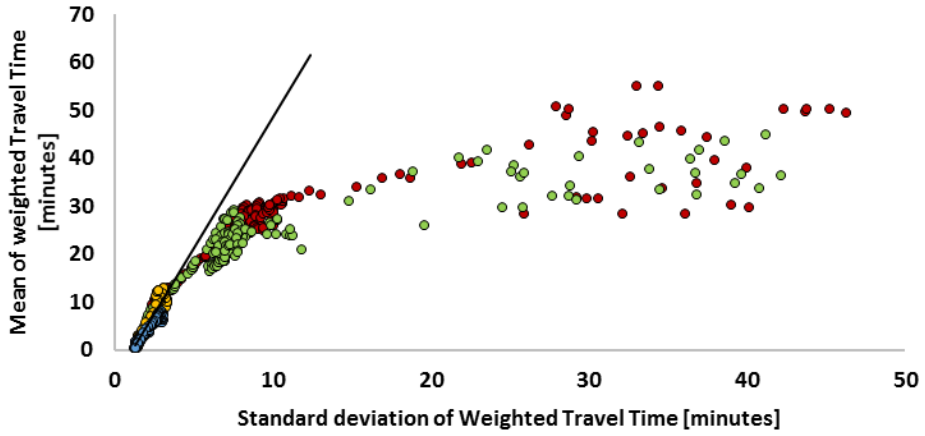
Results presented in Figure 4-9 demonstrate the effect connectivity has on the travel time reliability at low, medium, and high demand levels in a Salt Lake City testbed. At a low demand level (Figure 4-9a), travel time reliability increases with the increase in connectivity. As found in the Chicago

network case, the weighted mean and standard deviation of travel time at all connectivity levels follow a linear trend but to a different extent.

At the medium demand level (Figure 4-9b), travel time variability reduces significantly with the increase in connectivity. At 0% and 10% connectivity levels, high mean travel times and standard deviations are observed. When MPR reaches 50% or even 90%, the system does not exhibit high travel times and performs very reliably as compared to the other tested lower MPR cases. Hence, travel time variability reduces, and the system becomes more reliable with the increase in connectivity. At a high demand level (Figure 4-9c), the travel time variability follows the same trend at all connectivity levels, except for the MPR of 90%. The range of the mean travel time decreases with the increase in connectivity (i.e., high travel times are not observed with higher connectivity levels). At a 90% connectivity level, the range of mean travel times is as half as much as the range at lower connectivity levels. Hence, connectivity can improve the system's operational conditions by increasing throughput, lowering the maximum density attained, while reducing the travel time and its variability.


Mean of weighted Travel Time [minutes]
18
16
14
12
10
8
6
4
2
0
1
1.5
2
2.5
3
3.5
4
4.5
Standard deviation of Weighted Travel Time [minutes]

(a)


Mean of weighted Travel Time [minutes]
70
60
50
40
30
20
10
0
0
10
20
30
40
50
Standard deviation of Weighted Travel Time [minutes]

(b)

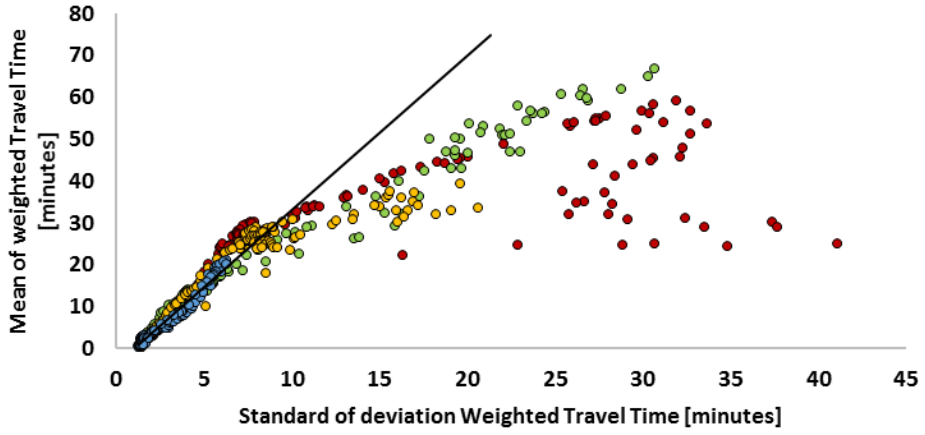

Mean of weighted Travel Time [minutes]
80
70
60
50
40
30
20
10
0
0
5
10
15
20
25
30
35
40
45
Standard of deviation Weighted Travel Time [minutes]

Figure 4-9 Travel time variability measures for Salt Lake City with (a) low demand. (b) medium demand, and (c) high demand.

4.2. Conclusion

The analysis presented in this chapter explores the network-level impact of Connected Vehicles technology on a transportation system's properties and travel time reliability. Traffic stream characteristics in a connected environment are identified based on a microscopic simulation tool. The calibrated speed-density relations are utilized in a mesoscopic simulation tool to study the network-wide effects of connectivity on travel time reliability. Simulations are conducted on two well calibrated, real-world networks, Chicago and Salt Lake City. An extensive calibration procedure of the networks and their respective traffic demands is out of the scope of this study, and the reader is encouraged to refer to (*144*).

This work confirms that the linear relationship between distance weighted travel time and distance weighted standard deviation holds at a network level and is not affected by either the demand level or the market penetration rate of connected vehicles. Hence, the network appears to retain its inherent properties (signature). Observations from the simulated traffic data show that with an increase in the MRP of connected vehicles, the network attains a lower maximum density and exhibits an increased flow rate for the same density level. Thus, a highly connected environment has the potential to help a congested network recover from flow breakdown and avoid gridlock. Moreover, the effects of connectivity become more prominent as the demand increases. Connectivity is found to be effective in improving the travel time reliability. Connected vehicles reduce the mean travel time while making the system more reliable. Overall, connectivity can improve system's performance by increasing throughput and travel time reliability at all demand levels.

CHAPTER 5. ADVANCED SIGNAL CONTROL STRATEGIES

Chapter 3 and Chapter 4 focused on the integration of microscopic and mesoscopic simulation environments. A successful integration enables an optimal trade-off between the numerical precision and computational efficiency. The integration was used to demonstrate the impact of connected vehicles on freeway and highway operations. If higher efficiency is to be achieved at a transportation system level, opportunities for improved performance need to be realized at both significant components, freeways, and arterials, concurrently. Higher throughput on freeways in mixed traffic would cause gridlock on the arterial street network unless a considerable improvement in signalized intersection control schemes is achieved (*147*). Hence, in this and the following two chapters, the focus is shifted to the arterial road network. Accordingly, signal control strategies for connected environments under different mixes of vehicular traffic are developed and tested in this part of the study. To emphasize, the focus of this effort is on the utilization of communication capabilities between vehicles and infrastructure to enhance the system's operational performance, and not on calibration of drivers' behavior models. .

State of the practice real-time signal control applications rely on infrastructure-based detection data, i.e., detector based adaptive signal control is dependent on detection accuracy and queue estimation. Once the queue grows beyond the length of the detector, or the link is over saturated, accurate measurements are not possible. Moreover, estimation models based on occupancy are used to estimate the flow, and it becomes difficult to differentiate between high flows, stopping, queue spillback, etc. With vehicle trajectories, more information is available, and this information is more accurate and freed of assumptions, imputation, estimation, or sophisticated statistical models. With technological advances, obtaining such detailed, more accurate and more reliable

information becomes achievable. Trajectory data from connected and automated vehicles offer more reliable real-time traffic information and represent an essential data source for a growing number of applications, including signal control strategies (*146*). Until recently its exploitation in the realm of adaptive traffic control has been limited. Naturally, better information enables the potential for better insight into operations, and consequently, improved control and management (*85*).

Previously proposed models used to optimize the traffic signal control might be complicated, and computationally expensive, sensitive to modeling errors. Furthermore, in order to facilitate the validation of the models, many scholars employed a simplified road or intersection model for simulation. However, road and intersections are relatively more complicated in reality, which poses a challenge to the adaptability of the models. This study attempts to address this gap by testing the strategy on a complex real-world network. The study addresses the question whether improvement is achievable, and if so, how significant it might be when only a fraction of vehicles is connected and automated. To this end, the proposed strategies compute vehicle-based performance metrics to optimize control parameters and individual vehicle trajectories.

In this study control strategies are being optimized to provide a higher quality of service, given the objective is chosen (minimizing delay, maximizing eco-friendly driving behavior, or multi-objective examining trade-offs, etc.) in a mixed traffic environment (automated vehicles, connected vehicles, and regular vehicles). More specifically, signals are optimized to combine the maximization of throughput with the minimization of delay, to keep vehicles moving, while clearing the queues. In the developed strategies, platoons of vehicles will not necessarily get service priority, and the control decision can be seen as the choice between clearing the queue or

progressing a platoon. The proposed advanced intersection-level control strategies are also designed to maintain the coordination along a corridor. Accordingly, CAVs are assumed to compute and transmit their delay accumulated along a route. The minimization problem accounts for the accumulated delay to enable traffic flow synchronization on the route. Traffic controllers are not communicating with each other, yet, V2X communication is presumed; controllers receive and process vehicle information as well as calculate and transmit signal information to the vehicles.

Flexible controller logic and phasing sequence/duration driven by vehicle-based computation of performance metrics distinguish this method from the ones found in the literature. Considering a simplified signalized intersection and roadway configuration, previous studies predominantly focused on control decisions of whether to switch or extend the active phase. An advanced traffic controller operates in an "intelligent fully-actuated" mode. Its setup accounts for various constraints: maximum/minimum green time or the maximum allowed waiting time, yet it enables phase skipping, extension, truncation, etc. when required.

Moreover, unlike previous work, this study incorporates the concept of marginal cost in its objective function calculation. When evaluating the next-optimal scenario, the proposed objective functions account for the negative externalities experienced by the vehicles on all approaches other than the one being served. This is to say that for any additional vehicle served, the delay experienced by vehicles on unserved approaches is augmented and balanced against the time savings gained, when deciding which phase should run next and for how long.

Low MPR of CAVs is one of the critical issues when relying on connected vehicle data, especially in signalized intersection-related applications. As a result, robustness, versatility, and effectiveness of the strategies are examined under different mixes of traffic. To this end, an intelligent fully-

actuated controller logic computes vehicle-based performance metrics to optimize control parameters and individual vehicle trajectories jointly.

5.1. Conceptual Framework

This study establishes a method to improve signal control strategies using individual vehicle trajectory information of a mixed fleet traffic environment. The proposed approach optimizes signal control parameters and trajectories of vehicles, interdependently and dynamically, to improve the traffic system's performance. Figure 5-1 shows the overall conceptual framework of the control strategies.

Three types of vehicles are considered in this study; **C**onnected (with connectivity but no automation), **A**utomated vehicles (with connectivity and automation), and **R**egular (with no connectivity and no automation).

Connected Vehicles (**C**Vs) are driven by human drivers and receive information from the surroundings through V2I and V2Vcommunication. Based on the information driver is provided with advice to control the vehicle. Final decisions regarding the vehicle's motion are at driver's discretion.

Automated Vehicles (**A**Vs) are the vehicles that are driven by robotic drivers and receive information from the surroundings through V2I and V2Vcommunication. All decisions are automated, and the vehicle is controlled accordingly. AVs can also represent a particular case of CVs where the driver has no reaction time and follows all conveyed messages for optimal flow.

Regular Vehicles (**R**Vs) are driven by human drivers and receive no information from the surroundings except what the driver can observe. Decisions regarding the vehicle's motion are based on the driver's perception of the conditions and the signage on the road facility.

In the study, different market penetration rates of **CAR**s are tested.

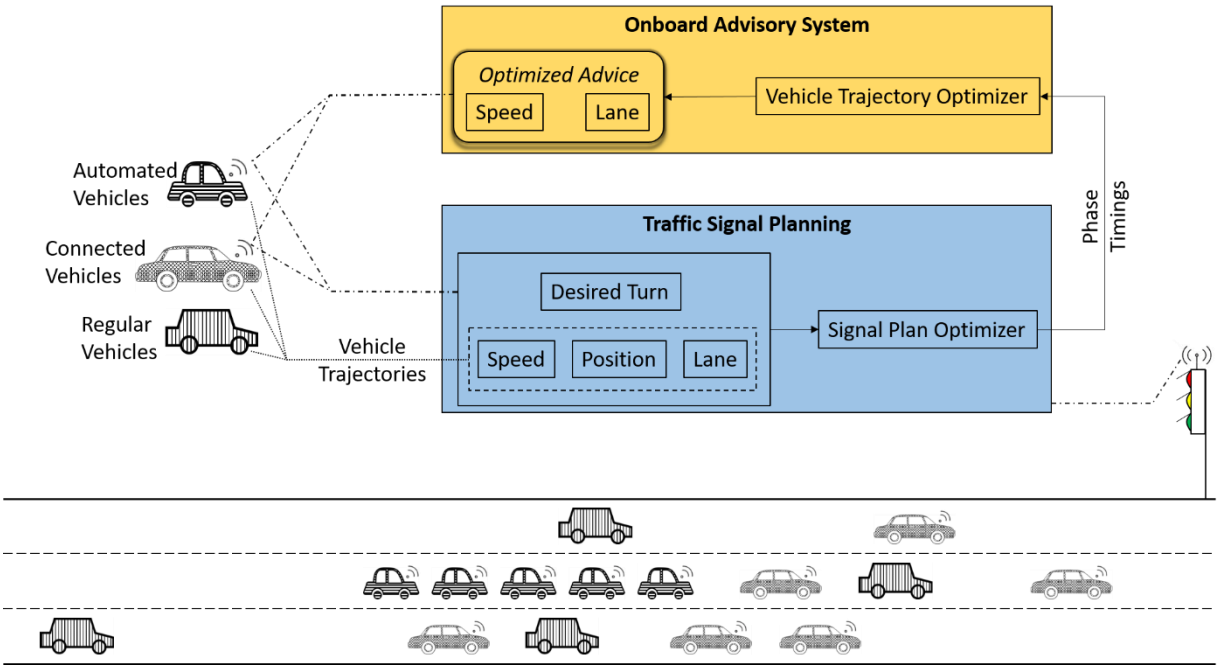


Figure 5-1 Schematics of the traffic signal control for mixed traffic environments

By assuming wireless communication, the CAVs can collect and disseminate traffic information and, the controller can provide meaningful data to the driver. CAVs periodically transmit information about themselves (every 1 second). Each record consists of a position, identification number, speed, direction, state, and a timestamp of the moment when the information was created. CAVs can receive optimized advice as to their speed and position.

Traffic controller can receive all the information CAVs are exchanging, thus finding out how crowded intersection approaches are. The conceptual framework assumes adaptive traffic controller logic. It is capable of optimizing control parameters based on the calculated traffic performance measures (Signal Plan Optimizer). Our control method is, thus, capable of accurately computing vehicle-based traffic metrics. The most important metrics used are individual vehicle delay and queue length.

Once the optimal signal timing scenario is known, the actual CAV's trajectory can be optimized. Vehicle Trajectory Optimizer, accounting for the queue length and optimal signal control parameters, determines the adequate vehicle behavior and sends optimized advice to all the CAVs.

The delay is calculated for each vehicle that passes through an intersection. The queue length and vehicles delay are computed by the traffic controller, who knows the traffic configuration every one second. For every vehicle record received, the controller checks it against its database. The record is stored and considered when calculating approach/movement parameters (demand, queue length, waiting time, etc.).

5.2. Simulating CAVs in VISSIM

To test and evaluate the methodologies simulation environment testbed was set up. The simulation framework presented here is primarily an evaluation tool of the method and strategies devised. Accordingly, a commercially available simulation environment VISSIM (*148*) was utilized. Please note that the methods presented in the dissertation are independent of the specific microscopic simulator used in this demonstration. Microscopic traffic simulation model of the real-world

corridor was used to emulate real-world roadway geometry and traffic conditions as well as the signal controller setup. In the prediction module, conventional vehicles are assumed to follow some standard car-following model, i.e., Wiedemann's. Wiedemann defines four modes of driving behavior: free driving, approaching, following, or braking. The car following behavior was modified for CAVs to represent enhanced operational features of these vehicles. Parameters such as look ahead distance, lane changing, reaction times, following headways, etc. were modified based on recommendations given in reference (*149*). Table 5-1 enumerates the parameters used in the Wiedemann's model for different vehicle types, i.e., ***CAR***s.

5.3. Overview of new control strategies

Basing on the conceptual framework, the next two chapters design and discuss two novel traffic control schemes. The first one represents a prediction-based adaptive traffic control strategy, while the second is a real-time platoon control strategy.

Prediction-based adaptive control method formulated in this research, in which vehicle trajectories provide the basis for predictive information to the signals on anticipated demand, forms the basis for timing plan settings optimization. Next, the information on upcoming signal changes provides input to vehicles to optimize their trajectories to reduce unnecessary idling and breaking.

Real-time platoon self-identifying control strategy represents a *platoon-phase scheduling heuristic* that considers clusters of vehicles as critical jobs. The framework devises an advanced, online, signal control logic for mixed traffic environments utilizing the information from CAVs to

augment controller/sensor data. A prerequisite of such an approach is the application of the innovative procedure for segmentation of traffic flows based on CV trajectory data.

Table 5-1 Car Following model parameters – Wiedemann 74 Model.

Parameter	Connected Vehicles	Automated Vehicles	Regular Vehicles
Average standstill distance (ft)	5.74	4.92	6.56
Additive part of safety distance	2.5	3	2
Multiplicative part of safety distance	1.875	2.25	1.5
Look ahead distance	0 to 1230 feet	0 to 1640 feet	0 to 820 feet
Look back distance	0 to 735 feet	0 to 980 feet	0 to 490 feet
Observed vehicles	6	10	2
Maximum deceleration - own vehicle (ft/s2)	-13.12	-13.12	-13.12
Maximum deceleration - trailing vehicle (ft/s2)	-9.84	-9.84	-9.84
-1 ft/s2 per distance - own vehicle and trailing vehicle (ft)	200	200	200
Accepted deceleration - own vehicle (ft/s2)	-3.28	-3.28	-3.28
Accepted deceleration - trailing vehicle (ft/s2)	-1.64	-1.64	-1.64
Minimum headway -front/rear (ft)	1.435	1.23	1.64
Safety distance reduction factor	0.525	0.45	0.6
Maximum deceleration for cooperative braking (ft/s2)	-11.48	-13.12	-9.84
Maximum speed difference (mph)	6.71	6.71	6.71
Maximum collision time (seconds)	10	10	10
Collision time gain (seconds)	2	2	2
Minimum longitudinal speed (mph)	2.24	2.24	2.24
Time before direction changes (seconds)	0	0	0
Overtake same lane vehicle -minimum lateral distance standing (ft)	0.5775	0.495	0.66
Overtake same lane vehicle - minimum lateral distance driving (ft)	2.87	2.46	3.28

To compare the novel - adjusted spatial longitudinal variation (ASLV) - clustering technique against a more conventional approach – a critical headway-based platooning - was also examined within the same control logic. The ASLV method differs from what transportation research and practice consider platooning. Reported results demonstrate that at corridor level, conventional gap-out platoon-based control, unlike the ASLV self-identification control method, fails to consistently achieve superior operational efficiency compared to the vehicle-actuated type of control.

Connected and automated vehicles communicate with the signals to convey their desired turning movement and accumulated delay along the corridor. This allows for the isolated controller to operate in an intelligent, yet, fully-actuated manner, recognizing the need to coordinate major direction traffic flows, i.e., to enable progression along the corridor, when warranted. In the next two chapters, the strategies are discussed in detail, separately.

CHAPTER 6. PREDICTION-BASED ADAPTIVE SIGNAL CONTROL

Advances in wireless communication have offered new possibilities for Intelligent Transportation Systems (ITS), aimed at improving driving safety and traffic efficiency. By adding short-range wireless communication capabilities to vehicles, these form a mobile network of vehicles that can exchange information with the infrastructure and among themselves. This study examines the possibility of deploying an adaptive signal control system at intersections, a system that can base its control decision on information coming from vehicles.

The effort presented in the chapter is aimed at designing a signal control strategy which reduces delay and offers fairness of progression to all the drivers in a mixed vehicular traffic environment using vehicle trajectories' information. Simultaneously, vehicle behavior is regulated with respect to the optimal signal scenario applied. The distinct contribution of the method is its ability to account for and optimize both components; i.e., it does not improve one aspect of the system at the expense of the other. While an improved control strategy is expected to enhance traffic performance, vehicle behavior is further adjusted, in response to signal indication encountered, to reduce idling thus serving an eco-friendly objective. Therefore, the proposed approach interdependently and dynamically determines signal timings and adjusts trajectories of vehicles to improve the traffic system's performance. Furthermore, unlike previous integrated solutions, which independently optimize certain signal plan operational features (either splits or cycle lengths or offsets), the method proposes to advance the deployed signal control strategy in real-time fully.

6.1. Predictive Traffic Control Strategy Framework

The overall framework comprises two interactive components: 1) real-time vehicle trajectories optimization and 2) predictive signal control parameters optimization. The methodological framework is presented in Figure 6-1. The framework defines the interaction between the two modules: predictive and real-time module. The selected corridor is modeled in the predictive module: this includes setting up a general intersection (corridor) configuration, demand levels, signal control parameters, etc. Based on prevailing demands from the field, the predictive unit evaluates a set of feasible solutions to determine the signal control scenario that optimizes the selected objective function. Signal control scenario which gives the optimal objective function value is identified as the next scenario and is communicated to the real-world controller. Simultaneously, CAVs also receive relevant information such as indication (and its duration) encountered at the first upstream signal. Real-world deploys an optimized signal timing strategy over the next ***Tplan*** seconds. While operational in the real world, an ongoing prediction decides the optimal strategy over the next ***Tplan*** seconds.

Real-world signal timing plans are known over the next ***Tplan*** period. Based on the signal timing plan, trajectories of CAVs are optimized to achieve an environmentally friendly objective. Trajectory uncertainty is eliminated for CAVs since they are not surprised by the end of green indication, unlike regular vehicles. Furthermore, the queue length in each lane of every approach is known.

At this stage, trajectory optimization takes place. CAVs adapt their speeds accordingly; based on when the current phase will end and how many cars are already queued CAVs will either avoid

unnecessary accelerations or react faster on the green. Fuel consumption and pollutant emissions are thus reduced.

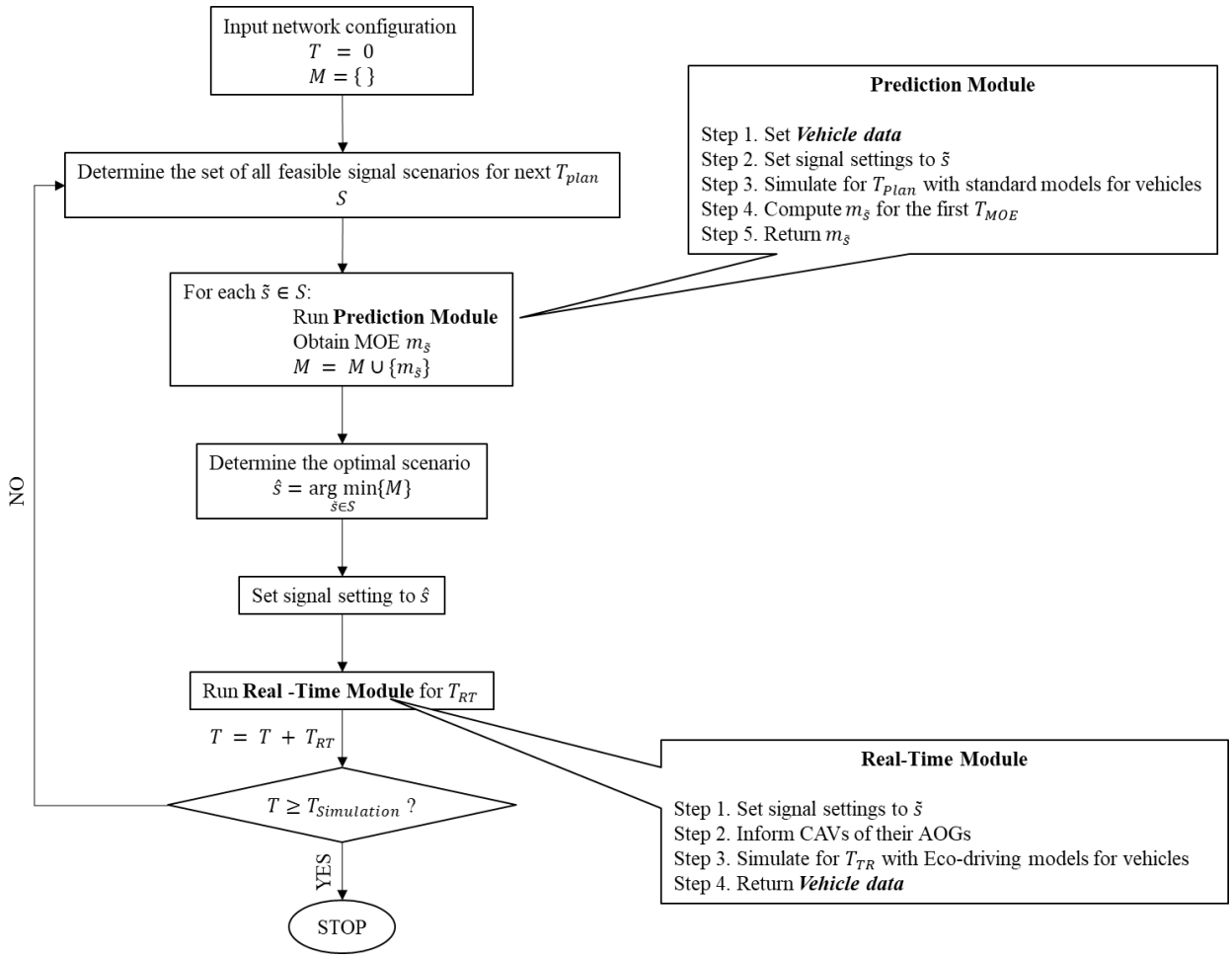


Figure 6-1 Modelling framework.

6.1.1. Prediction Signal Control Optimization (P-SCO)

P-SCO aims to dynamically optimize signal control settings, given the real-time traffic information. The solution method is based on a greedy heuristic (Figure 6-2) that has been designed

to optimize signal controller's phasing, sequence, and duration in an acyclic manner (over a period i.e., prediction horizon). A fully-actuated intelligent controller logic is presented. The controller keeps track of the vehicles on all approaches within the predefined distance, and accurately measures volume, queue length, and delay. Individual vehicle delay is computed, aggregated, and used as the basis of the optimization procedure. The following optimization problem is the core of P-SCO.

Given:

Φ: Set of feasible phasing scenario

Vehicle Data

$Veh\_Approach$: Set of vehicles approaching the intersection

$Delay_v, \forall\ v \in Veh\_Approach$

Parameters:

$\boldsymbol{m_{\tilde{s}}}$*: performance index of scenario* $\tilde{\boldsymbol{s}}$

$veh\_served_{\tilde{s}}$*: Set of vehicles who get served*

$veh\_unserved_{\tilde{s}}$*: Set of vehicles who do not get served*

Objective:

$$\underset{s \in \Phi}{arg\ min}\ \boldsymbol{m_{\tilde{s}}}$$

Such that:

Vehicles respect the laws of physics

Drivers follow the Wiedemann 74 car following model

No frequent transitions between signal plans so as not to disrupt the progression

Vehicles on the unserved vehicles are not waiting for more than $\boldsymbol{T_{Max}}$

$veh\_served_{\tilde{s}}$: vehicles in $Veh\_Approach$ who cross the intersection

$veh\_unserved_{\tilde{s}}$: vehicles in $Veh\_Approach$ who do not cross the intersection

$$\boldsymbol{m_{\tilde{s}}} = \left(\sum_{v \in veh\_unserved_{\tilde{s}}} Delay_v \right) \Big/ \left(\sum_{v' \in veh\_served_{\tilde{s}}} Delay_{v'} \right) \qquad (6\text{-}1)$$

Timing plan generation process takes place at the end of each phase and establishes a plan for the following ***Tplan*** duration based on the measured parameters. Further phasing adjustments may occur, such as phase skipping, extension, rotation, and truncation. Controller logic accounts for various limitations, such as max/min Green time or the maximum allowed waiting time, also, gap and max out are recognized. There is no notion of cycle length. Phasing is performance metrics - driven, i.e., delay of still unserved vehicles vs. total delay of served vehicles. Note that *unusual waiting* times on any approach, are considered a priority when establishing a phasing sequence and these vehicles will be served first.

The algorithm determines the optimal control settings to service the predicted demand by minimizing the objective function, $\boldsymbol{m_{\tilde{s}}}$. The objective function considers an individual vehicle's delay, the negative externality of serving an approach and number of served and unserved vehicles. The objective function addresses two conflicting aspects: minimizing total delay while maximizing

throughput, it is calculated as the total vehicle delay on all unserved approaches divided by the total delay experienced by the vehicles being served. However, the choice of the objective function can be considered planner's preference.

At this stage, vehicle progression under different signal phase timings is predicted and utilized to determine the most desirable signal scenario with respect to prevailing traffic conditions. A microscopic traffic simulation model of the real-world corridor is the core of the prediction module. It is intended to emulate real-world roadway geometry and traffic conditions as well as the signal controller setup. In the prediction module, conventional vehicles are assumed to follow some standard car-following model, e.g., Wiedemann's. Wiedemann defines four modes of driving behavior: free driving, approaching, following, or braking. The car following behavior was modified for CAVs to represent enhanced operational features of these vehicles. Parameters such as look ahead distance, lane changing, reaction times, following headways, etc. were modified based on recommendations given in reference (Stanek et al., 2018).

Given current (real-time) signal settings and ***vehicle data***, all possible signal timing/phasing settings are identified to form the vector, $\boldsymbol{S}$. "Feasible phasing scenario" refers to potential phase combinations that account for and give the right of way to the movement if: (1) the number of inbound vehicles at an intersection approach/movement is the highest compared to that of any other approach and (2) there is no vehicle at any other approach that had been waiting longer than a pre-set maximum time $\boldsymbol{T_{Max}}$ (ex. 120sec). Each scenario that satisfies these two conditions serves the critical inbound demand and becomes a part of $\boldsymbol{S}$. Vector of scenarios, $\boldsymbol{S}$, evaluated when identifying the optimal scenario. For example, in a standard NEMA-RBC controller setting, an

approach corresponding to EBL (phase 1) has the longest queue, so either 1 and 5 or 1 and 6 can be implemented to serve this demand. (*150*) These two combinations will be included in $\boldsymbol{S}$. Here the phase combinations consist of any two non-conflicting signal groups. To determine which phase should run next and for how long $\boldsymbol{S}$ consists of higher ranked phasing scenarios (in terms of greater queue length or waiting time). The purpose of $\boldsymbol{S}$ is to limit the search space and reduce computational effort.

Then each of the feasible signal timing settings, $\tilde{\boldsymbol{s}} \in \boldsymbol{S}$, is tested, one by one. Under each $\tilde{\boldsymbol{s}}$ scenario, corresponding objective function value, $\boldsymbol{m}_{\tilde{\boldsymbol{s}}}$, is recorded and a vector $\boldsymbol{M}$, ($\boldsymbol{m}_{\tilde{\boldsymbol{s}}} \in \boldsymbol{M}$) is created. Signal control settings associated with the optimal objective function value is determined and implemented in real-time. If more than one of the scenarios result in an optimal objective function value, signal timing plan most similar to the current one is preferred.

In the predictive module, for each $\tilde{\boldsymbol{s}} \in \boldsymbol{S}$ at every time step, the delay of served and unserved vehicles is being computed. In addition, maximum queue is recorded, regardless of the approach, as well as vehicle's earliest entry on any approach, other than the one being served. Low demand conditions allow for current phase duration to be extended if no vehicle is waiting to be serviced at any other approach. High demand conditions require a phase maximum time to be defined. Earliest entry on a conflicting (or any other) approach starts the "*Max Green*" timer for an active phase.

The time it takes to clear the queued vehicles at an approach/movement represents the optimal duration of a phase -$Duration_{\tilde{\boldsymbol{s}}}$. Minimum $\boldsymbol{m}_{\tilde{\boldsymbol{s}}}$ phase scenario will become the $NextPhase$ to run in $\boldsymbol{RT}$ and a part of the optimal signal information fed to $\boldsymbol{RT}$. Once $NextPhase$ is determined, $\boldsymbol{S}$

is updated. Previous step's $NextPhase$ will not be included in the new $\boldsymbol{S}$ unless extended due to lack of demand on conflicting approaches/movements.

Optimal adaptive traffic signal determining algorithm

Require: snapshot of current traffic condition, vehicles' accumulated delay and desired turns

Objective: Minimize the delay at the intersection and externality of serving approach

Procedure: Efficient Signal timing plan

for $\tilde{\boldsymbol{s}}$ in $\boldsymbol{S}$:

$Delay_{served}$ = 0

$Delay_{unserved}$ = 0

$Duration_{\tilde{\boldsymbol{s}}}$ = 0

while ($Wait_{vehicles} < T_{max}$ and $Demand_{PS} > 0$):

$$Delay_{served} \mathrel{+}= \sum_{v \in veh\_\tilde{s}\_approac\ h} Delay_v$$

$$Delay_{unserved} \mathrel{+}= \sum_{\substack{v \in veh\_\widetilde{not-s}\_approac\ h \\ \widetilde{not-s} \in \boldsymbol{S} \setminus \tilde{\boldsymbol{s}}}} Delay_v$$

// $Delay_v$ is control delay for all regular vehicles and delay accumulated over the corridor for the connected or automated vehicles.

$$Duration_{\tilde{\boldsymbol{s}}} \mathrel{+}= TimeStep$$

$$\boldsymbol{m}_{\tilde{\boldsymbol{s}}} = \frac{Delay_{unserved}}{Delay_{served}}$$

$$NextPhase = \arg\min_{\tilde{\boldsymbol{s}} \in \boldsymbol{S}} \boldsymbol{m}_{\tilde{\boldsymbol{s}}}$$

$$Duration = Duration_{NextPhase}$$

Figure 6-2. Predictive adaptive signal control algorithm

The procedure is repeated until the entire prediction period is evaluated and as a result, optimal phasing order (and respective durations) are known over the next $\boldsymbol{T_{Plan}}$ duration.

6.1.2. Real-time Component

The real-time stage represents the actual real-world timeline and a real-world signalized corridor where the advanced signal control strategy is to be implemented. As mentioned above, the information exchange between a connected vehicle and the signal controller is enabled. The traffic signal system can retrieve high resolution (frequency of 1 second or higher) phase status and duration data as well as vehicle information within a pre-specified distance from the controller. Furthermore, the controller exchanges relevant information with the prediction component. It is also capable of computing traffic performance metrics through wireless communication of the traffic signal system with vehicles. Individual vehicle delay and queue length are considered most relevant in this analysis. The queue length and control delay is computed by the traffic controller, which is aware of the traffic configuration at every time step. An essential feature of this work is that CAVs are capable of computing their delays, accumulated along the corridor, and transmit it to the controller. Accumulated delay forms the basis of the decentralized strategy of flow synchronization over a facility. At this stage, CAVs are assumed to drive in an eco-friendly manner. Hence, their optimized speeds are determined.

The heuristic in Figure 6-3 was designed to achieve vehicle trajectory adjustments with respect to anticipated green/red time onset and duration. Signal timing information was considered essential since no real-time trajectory adjustment can be realized without knowing signal control parameters a priori. Inbound vehicles are only considered, and once the stop bar is crossed, they are disregarded from signal/trajectory-related calculations.

Speed Advisory heuristics for adaptive traffic signal

Require: Downstream signal plan, vehicles' current progression status, vehicle attributes

Objective: Reduce idling at the intersection

Procedure: Efficient speed advisory

for veh_i in $All\_Vehicles$:

if $veh_i \in \{CV, AV\}$:

Compute $Dist_i$; Distance of veh_i to the downstream signal

Compute V_{Green} and V_{Red} for the veh_i,

// V_Φ is the speed that veh_i needs to arrive at the start of the next Φ phase. If the current phase is already Φ, V_Φ is set to a high number.

if ($V_{Green} < V_{prevailing}$):

$V_{desire} = V_{Green}$

else if ($V_{Red} > V_{prevailing}$):

$V_{desire} = V_{Min}$

else if ($V_{Desire} \cong V_{prevailing}$ and $V_{Red} < V_{prevailing}$):

$V_{desire} = V_{SpeedLimit}$

Figure 6-3 Real-time speed advisory algorithm.

In this study, queue lengths and arrivals on green are determined within the predictive module. These are calculated as the two parameters in the predictive stage and transmitted to the controller in real-time. This property illustrates another novelty of the proposed approach. The method transfers state variables, dynamically, from predictive to the real-time component of the system. As soon as this information is known in real-time, the associated module, on a second-by-second basis, adjusts CAV's speeds to reduce idling along the trajectory. Based on the speed advisory heuristic presented in Figure 6-3, CAVs are being advised to accelerate if it is reasonable to cross

the intersection given the information regarding the next green indication and corresponding duration. Distance to stop bar and required range of speeds (V_{Red}, V_{Green}) is computed for each connected/automated vehicle. A vehicle needs to attain the speed V_{Red} in order to reach the stop bar at the next onset of red; similarly, speed V_{Green} guarantees that the vehicle will clear the intersection within the next green indication. If the vehicle is moving at a speed lower than V_{Green}, its desired speed is set to be V_{Green}. If the vehicle is anticipated to arrive at the beginning of next red, its desired speed is reduced to its minimal speed V_{Min} to avoid idling i.e., stopping. In case current prevailing speed is, approximately, the desired speed, yet higher than V_{Red}, the desired speed is set to be the maximum allowed speed i.e. the speed limit.

6.2. TESTBED SETUP AND IMPLEMENTATION

A simulation environment testbed set up is required to test and evaluate the methodology. Two identical microsimulation models were utilized to investigate the validity of the proposed conceptual framework: (1) Real-Time (***RT***) module, emulating real world and running in real time, and (2) Prediction (***P***) module, running as virtual reality, predicting traffic conditions. The ***RT*** and ***P*** modules are being run simultaneously; the ***P*** module designs an advanced signal control strategy, while the ***RT*** is implementing it and optimizing vehicle trajectories in response to said optimal strategy in real time. In this regard, the framework adopts a similar architecture as previous traffic estimation and prediction systems (*151*).

Based on predicted trajectories, optimization of signal control settings is performed, and at the end of each prediction horizon, optimal signal parameters are fed to the real world. The adaptive

control method performs a rolling horizon-based optimization of signal settings via second/parallel simulation. The adaptive traffic control design assumes the controller, within the ***P*** module, decides the phasing according to the traffic configuration. Prediction uncertainty is modeled through a compliance rate of CAVs, acknowledging only the rate itself is unchanged when predicting trajectories. Vehicle trajectories are predicted based on the current field conditions and are expected to differ from the ones in real time. Furthermore, the two modules: ***RT*** and ***P*** used different random seeds.

6.2.1. Communication between modules

The two modules (models) communicate by transferring vehicle and signal information at predetermined time intervals. Real-Time Horizon represents a roll period when vehicle information is transmitted. Prediction Horizon represents the duration of time for which advanced control strategy is designed.

At the beginning of each Prediction period, vehicle and signal controller information, representing the current state of the system in real time, are captured in ***RT*** and transferred to ***P***. They contain the data related to all the vehicles in the network (location and position on approaches/movements) as well as signal timing – phasing and duration.

At the end of each Prediction period, the controller retrieves predicted signal timing information - a sequence of phases along with associated durations - to be implemented in real time, over the next prediction horizon. It is essential to recognize the existence of the overlap (simulation running) time between the two modules.

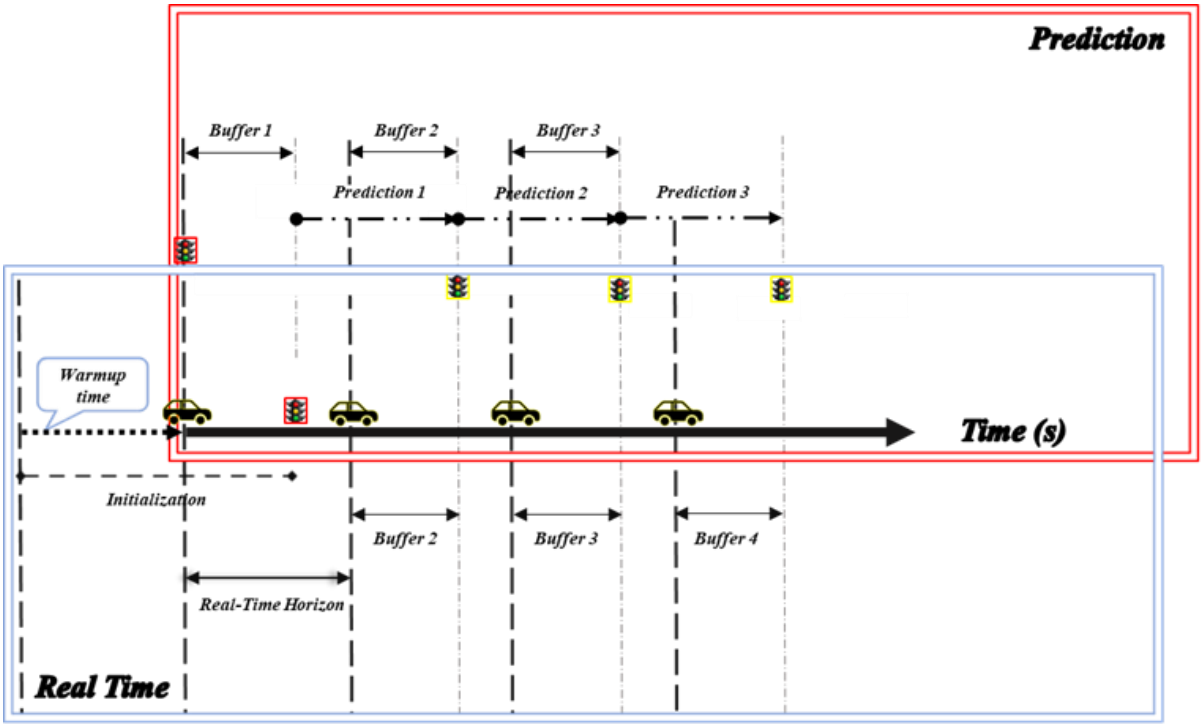


Figure 6-4 Interaction between the modules.

6.2.2. Overlap (BUFFER) time

After initialization, buffer time represents the overlap in a simulation running time between the two modules. This overlap period initializes each prediction cycle. Each Real-Time Horizon ends, and each Prediction initializes representing the same signal and vehicle system state. At the start of each overlap time, vehicle state information is transferred from ***RT*** to ***P***.

Since *overlap* related signal information is already available in ***P*** - from the previous prediction - once vehicle information is exchanged between the two modules a new prediction cycle can start: first reproducing overlap, then the actual prediction time.

Buffer time serves three purposes: 1) eliminates abrupt switching between phases when going from ***RT*** to ***P*** and vice versa, 2) ensures a continuous and seamless information exchange between the

modules, 3) enables ***P*** to "warm-up" by incorporating previous signal and vehicle states before phasing scenarios evaluation begins.

6.2.3. Real-time Module

The real-time simulation model emulates the real-world traffic network and its operational characteristics. The simulation model runs at the same rate as the actual physical system.

A warm-up period of 300 seconds is assumed, of which the last period representing the buffer time (warmup overlap). Initial signal control parameters obtained from the warmup overlap are fed to prediction.

At the beginning of each Real-Time Horizon, vehicle information is being communicated from ***RT*** to ***P*** and traffic conditions prediction is being executed. Real-Time Horizon assumed here is 180 seconds. At the end of Prediction Horizon, ***P*** module has generated an optimized signal control strategy, and the ***RT*** module implements it in real-time. These two data exchanges are executed continuously in real time.

6.2.4. Prediction Module

The interaction between the two modules, ***RT P***, is continuous and occurs at predetermined intervals. It should be noted that in addition to (before the beginning of any) Prediction Horizon time, overlap time is also being simulated in ***P***. This time is not considered as part of the Prediction Horizon time.

The ***P*** module replicates the overlap time sequence of phases and their duration while predicting traffic conditions based on the "current state of the system." The current state of the system characterizes real-world traffic conditions as if an aerial snapshot was taken referencing actual locations and positions of the vehicles along the corridor.

The current ***RT*** signal controller state and timing information from the previous prediction, i.e., ***P*** cycle is thus equivalent and, as soon as current vehicle data from ***RT*** is available, the new prediction cycle can begin. Once this overlap time runs out, prediction evaluates which signal group should run next and for how long. When these are determined, ***P*** module re-simulates *overlap* time and continues onto the signal group sequence identified as optimal in the previous steps, so that traffic demand and vehicle interactions can be re-estimated iteratively Figure 6-4. The procedure is repeated until the entire Prediction Horizon is exhausted and therefore optimal control strategy determined. At this point, the ***Tplan*** is communicated to (and applied in) ***RT*** over the next 180 seconds (Real Time horizon). ().

6.3. Vehicle-Trajectory based Signal Control Optimization Logic

The method to dynamically optimize signal control settings is applied within the predictive module.

An algorithm has been designed to optimize the signal controller's phasing, sequence, and duration in an acyclic manner (over a period i.e., Prediction Horizon). The heuristic determines the optimal control strategy to service the predicted demand by minimizing the objective function. For each feasible phase scenario (***PS***) objective function value is being evaluated at each time step. *The*

objective function considers two aspects: minimizing total delay while maximizing throughput. For a chosen phasing scenario, it is calculated as the total delay on all unserved approaches divided by the total delay of served vehicles; please refer to the equation (6-1).

"Feasible phasing scenario" refers to the potential phase combinations that account for and give the right of way to the movement if

1) the number of inbound vehicles on an intersection approach/movement is the highest compared to that of any other approach, and

2) there is no vehicle at any other approach that had been waiting longer than a pre-set maximum time, i.e. ***T_max*** (e.g., 120sec).

Each ***PS*** that satisfies these two conditions serves the critical inbound demand and becomes a part of the CASELIST. CASELIST is a vector of scenarios that will be considered for evaluation when identifying the optimal scenario. For example, in a standard NEMA-RBC controller setting, an approach corresponding to EBL (e.g., phase 1) has the longest queue, so either 1 and 5 or 1 and six can service this demand (*150*). Accordingly, these two ***PS***s are included in the CASELIST. Here, the phase combinations consist of any two non-conflicting signal groups.

CASELIST consists of higher ranked ***PS***s (regarding greater queue length or waiting time) when determining which phase should run next and for how long. The purpose of the CASELIST is to limit the search space and reduce computational effort.

In the ***P*** module, under current ***PS***, at every time step, the number of Queued (***Q***) and Served (***S***) vehicles is being recorded. ***Q*** and ***S*** correspond to the movement being served under evaluated ***PS***. Also, maximum Queue (***max_Q***) is recorded, as well as vehicle's earliest entry on any approach,

other than the one being served. The last two parameters are important when accounting for demand extremes. Low demand conditions allow for current ***PS*** duration to be extended if no vehicle is waiting to be serviced at any other approach. High demand conditions require a maximum phase time to be defined. Earliest entry on a conflicting (or any other) approach starts the maxGreen timer for an active phase.

The time it takes to clear the queued vehicles at an approach/movement represents the optimal duration of a phase scenario under evaluation. For each ***PS*** in the CASELIST, optimal duration and objective function value is being calculated. It will become a part of the optimal signal file fed to ***RT*** if the associated objective function value is found to be optimal. The optimal objective function value will determine which ***PS*** will be applied next in the ***RT***.

Once the next optimal phasing scenario to execute in ***RT*** is determined, the CASELIST is updated, once again enumerating only feasible ***PS***s. Previous step's optimal ***PS*** will not be included in the new CASELIST unless extended due to lack of demand on conflicting approaches/movements.

The procedure is being repeated until the entire prediction period is evaluated and as a result, optimal phasing order (and their durations) are known over the next 180 seconds. At this time optimal signal timings file is transferred to the ***RT*** module and starts running in real time.

6.3.1. Signal Controller Logic

The advanced intelligent fully actuated type of controller operation builds on the logic of a NEMA-RBC controller (*150*). Phase status and duration is determined based on the queue length and delay experienced by individual vehicles. The controller continuously monitors the vehicles within the

predefined distance from the controller, so it can measure both volumes and delay accurately. The queue-based delay is calculated and used as the basis of the optimization procedure.

The timing plan generation process takes place once during each period and establishes signal timing over the next ***Tplan*** seconds based on the measured parameters. However, further phasing adjustments may occur, such as phase skipping, extension, rotation, and truncation.

Acyclic controller nature is assumed – there is no notion of cycle length. Phasing is driven by computing the ratio: served vs. total unserved vehicles' delay.

The signal plan is adjusted to meet various limitations, such as max/min Green time or the maximum allowed waiting time. After the minimum green time, the phase *gaps out* if no (or less than a pre-specified number of) incoming vehicles is detected. It should be noted that the concept is considering a distance *gap out*. The extension is enabled if vehicles are still incoming, but no other movement is calling a phase. Additionally, if next (optimal scenario) phase is Green and determined to be less than minGreen, active phase is extended until an acceptable (minimum) number of vehicles are recorded on any of the approaches. Duration based *max out*, like with RBC controllers, accounts for earliest vehicle entry on any approach-movement except the one served. Phase Rotation/Skipping is only possible after a phase maxes out and no vehicle is waiting longer on any other approach than the one calling this phase. Low demand allows for optimal green time to be shortened.

Note that *unusual waiting* times on any approach, are considered a priority when establishing a phasing sequence and those vehicles will get served first.

6.4. Numerical Experiments Results

The method proposed was evaluated on two testbeds.

1. An isolated-intersection under different MPRs of CARs and different demand levels.
2. A real-world corridor under calibrated demand under different MPRs of CARs.

An isolated intersection in (1) is a part of the real-world corridor in (2) and the experiments are designed in such a way to first analyze the validity of the strategy on an intersection-control level. Next, to extend the concept and evaluate its potential to, de facto, synchronize traffic flow, when warranted by, CAVs accumulated delay, the analysis is conducted on a corridor level as well.

As shown in Table 6-1, in both cases, isolated intersection and corridor, 4 cases of traffic mixes, i.e., different MPRs of CARs, are considered.

Table 6-1 Summary of the MPRs of CARs.

MPR Case	Market Penetration Rate (%)		
	C*onnected Vehicles*	***A****utomated Vehicles*	***R****egular Vehicles*
1	0.0 %	0.0 %	100.00 %
2	33.33 %	33.33 %	33.33 %
3	0.0 %	100.00 %	0.0 %
4	100.00 %	0.0 %	0.0 %

East-West direction is assumed to be the major direction. As shown in Table 6-2, three demand levels: low, medium, and high, are considered for the isolated intersection. Demand scenarios were

designed to cover a 2-hour period, where each demand level is simulated for 30 minutes, starting from low, then medium, high and the simulation ends with a medium demand level. The chosen high demand level is unrealistically high and was chosen to test the capabilities of the strategy under extreme conditions. The corridor-level analysis considered real-world calibrated demand and the experiments were conducted to cover a 1-hour period.

Table 6-2 Flow on each approach under different demand levels.

Demand Case	**Direction Flow (vehicles/hour/lane)**			
	East	West	North	South
Low	600	600	270	270
Medium	800	800	360	360
High	1200	1200	540	540

Three measures of effectiveness (MOEs) are presented in Figure 6-5 through Figure 6-7, for each of the testbeds. Active signal-speed advisory strategy in a mixed traffic environment significantly improves traffic operations concerning primary performance indicators such as total delay, stopped delay and queue length. The results are aggregated for all approaches, over a single evaluation period, for each of the cases reported. As for the isolated intersection, overall queue lengths were reduced by as much as 74% (period 7 and 8, traffic mix 1 and 2). Comparing all of the approaches, total delay decreased by at most 33 %, while stopped delay was reduced by 43% at most.

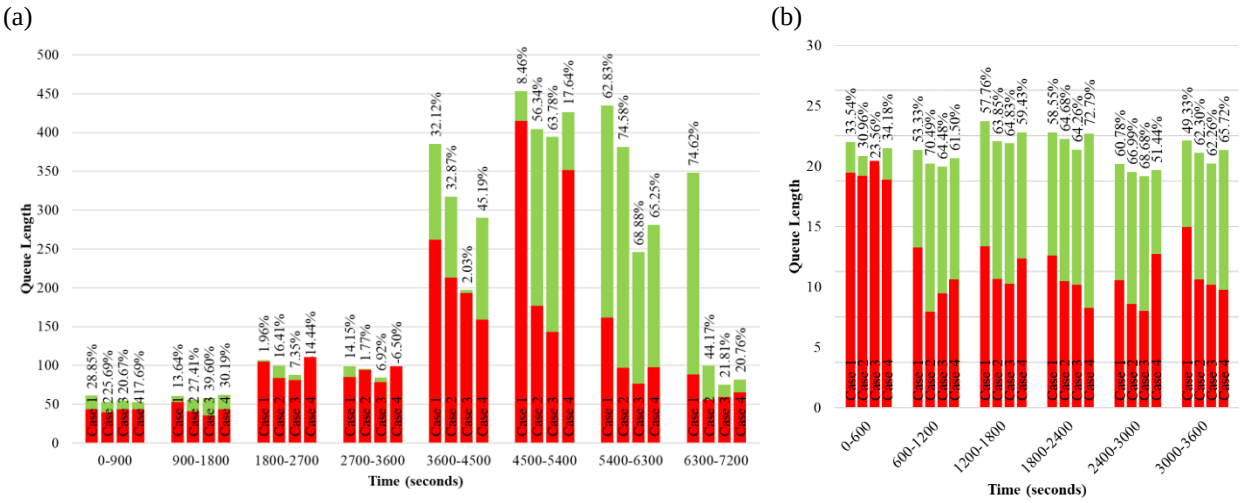


Figure 6-5. Queue length reduction (a) isolated intersection; (b) whole Corridor.

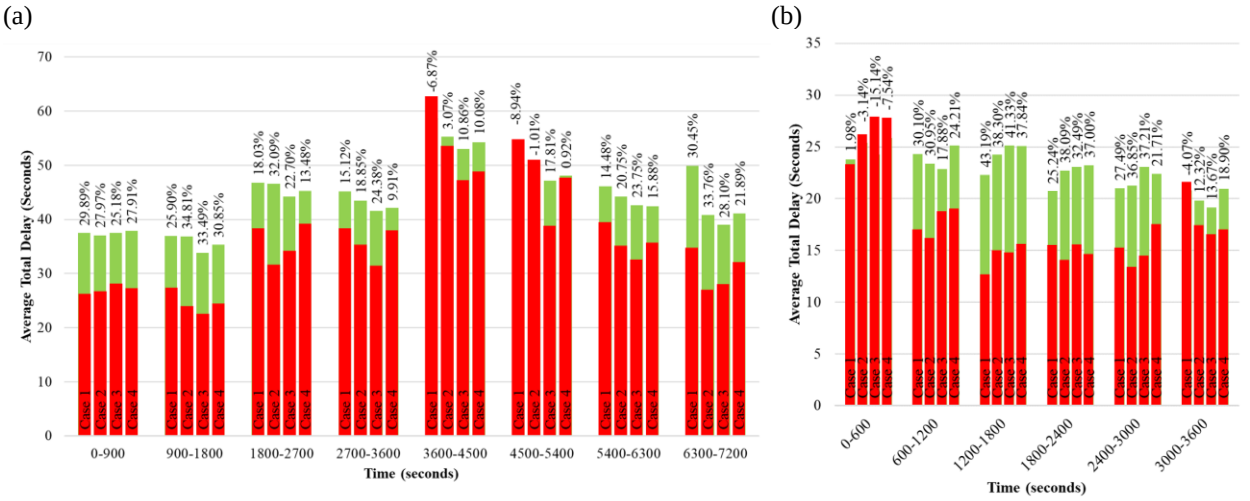


Figure 6-6. Total delay reduction (a) isolated intersection; (b) whole Corridor.

As for the overall corridor performance, the figures below establish a global, positive, trend as well, demonstrating considerable improvement in each of the three MOEs. Fuel consumption and emissions savings seemed modest, at best. Up to 10% reduction in fuel consumption was observed,

but the results were not consistent throughout the experiments. Conversely, overall delay stopped delay, and queue length was significantly reduced in all tested scenarios, with, in some cases, up to 43% (delay) and 66% (queue length) reduction.

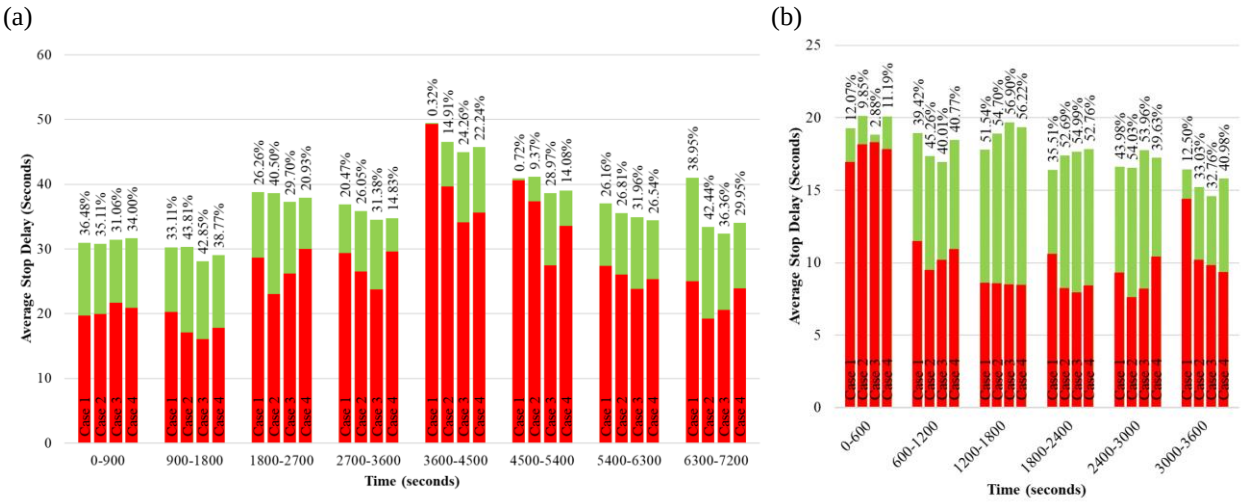


Figure 6-7. Stopped delay reduction (a) isolated intersection; (b) the whole Corridor.

6.5. Conclusion

The framework proposed in this study formulated an adaptive traffic signal system based on short-range wireless V2X communication. With traffic configuration fed continuously to the traffic controller, vehicle-based performance metrics were computed to optimize control parameters and vehicle speeds. CAVs were presumed to calculate and transmit their accumulated delay along a route. Accumulated delay formed the basis of the decentralized strategy of flow synchronization over the facility.

A microsimulation-based prediction module was designed to project the demand and optimize control parameters to be applied in real-time. In this setting, connected and automated vehicles communicate with the signals to convey their desired turning movement and accumulated delay along the corridor. Communication of such information allows for the isolated controller to operate in an intelligent, yet, fully-actuated manner, recognizing the need to coordinate major direction traffic flows, i.e., to enable progression along the corridor, when needed. Furthermore, there is no issue of privacy violation as vehicles are not being tracked throughout their route. Also, controllers do not store any data regarding the vehicles once they clear the intersection.

This architecture demonstrated clear benefits over current adaptive systems based on sensors or cameras. An integrated modeling framework to validate the method and its effectiveness in a microsimulation environment was developed. Testbeds included a corridor and an isolated intersection that is a part of a signalized urban arterial.

The focus of the analysis was on traffic operations during oversaturation, and how the system behaves in such circumstances. Total and stopped delays, as well as queue lengths, were significantly reduced with the strategy applied; however, pollutant emissions did not decrease consistently. Please note that the authors did not derive the fuel consumption measurements, these were the output of the microsimulation tool used. Of course, as more of the vehicle fleet becomes electrified, this particular performance metric becomes a less important consideration. The rest of the operational metrics were calculated independently of the tool.

Although the study puts forward an innovative and promising approach that provides a useful foundation for traffic control in a mixed connected vehicle traffic environment, several implementation challenges would have to be addressed, including:

- CAVs need to convey their information to the intersection controller infrastructure which means that data transmitted should be of a particular format. Hence, there is a need for the standardization of the format and the exchange process itself.
- CAVs, currently, do not estimate/calculate their delay. Technology to achieve this should be such that the users are not able to change the parameters and the models utilized are to be standardized.

CHAPTER 7. REAL-TIME PLATOON SELF-IDENTIFYING ADAPTIVE SIGNAL CONTROL

As established previously, with the emergence of connected vehicle technology, an abundant amount of high-resolution data become available. Trajectory data from connected/automated vehicles offer more reliable real-time traffic information and represent an essential data source for a growing number of applications, including signal control strategies (*146*). Until recently its exploitation in the realm of adaptive traffic control has been limited. State of the practice real-time signal control applications relies on infrastructure-based detection data. With technological advances, obtaining detailed, more accurate and more reliable information becomes achievable. Naturally, better information enables the potential for better insight into operations, and consequently, improved control and management (*85*).

If higher efficiency is to be achieved at a transportation system level, opportunities for improved performance need to be realized at both significant components, freeways, and arterials, concurrently. Higher throughput on freeways in mixed traffic would cause gridlock on the arterial street network unless a considerable improvement in signalized intersection control schemes is achieved (*147*).

Consequently, this study combines the maximization of throughput with the minimization of delay, to keep vehicles moving, while clearing the queues, whenever warranted by the delay incurred. The developed strategies will not necessarily give service priority to the largest platoon as the control decision is tradeoffs between clearing the queue and progressing a platoon.

The proposed control scheme represents a *platoon-phase scheduling heuristic* that considers clusters of vehicles as critical jobs. The basic premise of the proposed control strategy is that a platoon should be given preferential treatment, so it can traverse the intersection unimpeded by a red signal indication when delay savings are more substantial compared to any of the alternative phase-platoon pairs. This framework formulates an advanced, real-time, signal control logic for mixed traffic environments utilizing the information from CAVs to augment controller/sensor data. A prerequisite of such an approach is the application of the formulated procedure for segmentation of traffic flows based on CV trajectory data.

This study incorporates vehicle-following behavior standards into its platoon self-identification logic. The drivers are expected to adjust their behavior according to the actions of their immediate surroundings, i.e., other drivers. Consequently, the vehicle's actual position is dependent on whether a preceding vehicle can be identified, and if so, acknowledging longitudinal spatial adjustment applies.

This study proposes an advanced decentralized control strategy, designed to also when warranted, maintain the coordination along the corridor. Accordingly, CAVs are, as was already established in the previous chapter, assumed to compute and transmit their accumulated delay, along a route, thus enabling corridor traffic flow synchronization when incurred delay justifies such actions. Moreover, this study incorporates the concept of marginal cost in its objective function calculation. When evaluating the next-optimal scenario, the proposed objective function accounts for the negative externalities imposed by the said scenario.

Flexible controller logic and phasing sequence/duration driven by vehicle-based computation of performance metrics distinguish this method and one presented in Chapter 6 from the ones found

in the literature. Considering the overly simplified signalized intersection, signal, and roadway configurations, previous studies predominantly focused on control decisions whether to switch or extend the active phase to accommodate an incoming platoon. As mentioned in Chapter 6, an advanced traffic controller operates in an "intelligent fully-actuated" mode. Its setup accounts for various constraints: maximum/minimum green time or maximum allowed waiting time, yet it enables phase skipping, extension, rotation, truncation, etc., when required. Connected and automated vehicles communicate with the signals to convey their desired turning movement and accumulated delay along the corridor. This allows for the isolated controller to operate in an intelligent, yet, fully-actuated manner, recognizing the need to coordinate major direction traffic flows, i.e., to enable progression along the corridor, when warranted.

Differing from Prediction-based adaptive control (Chapter 6), this chapter presents real-time platoon self-identifying control strategy. The strategy represents a *platoon-phase scheduling heuristic* that considers clusters of vehicles as critical jobs. The framework devises an advanced, online, signal control logic for mixed traffic environments utilizing the information from CAVs to augment controller/sensor data. A prerequisite of such an approach is the application of the innovative procedure for segmentation of traffic flows based on CV trajectory data.

The generic structure of the control algorithm is intended to handle a variety of traffic conditions successfully such as, undersaturated and oversaturated flow conditions. Furthermore, the formal framework introduced in this study is novel and generic such that it has the potential to be applied to any intersection/signal configuration.

Low penetration rate is one of the critical issues with relying on connected vehicle data, especially in signalized intersection-related applications. The novelty of the proposed approach along with

connected vehicles' data availability limits its design, application, and testing to a simulation-based environment. The traffic controller can record intersection traffic state, i.e., vehicle trajectories data and signal events via screenshot-type of feature and process this information in real-time to determine the next best phase-platoon pair. Furthermore, marginal cost-based phase-platoon allocation algorithm is proposed to optimize phase scheduling in real-time. By tracking traffic configuration on a second-by-second basis, the controller determines the actual green time duration in real-time.

7.1. Platoon Identification and Phase Allocation Control Method

This section puts forward a real-time signal control strategy that uses vehicle trajectories information to decide which phase-platoon should be given the right of way. Phase-platoon control strategy is based on a procedure referred to as platoon self-identification – a novel approach to traffic flow partitioning named adjusted spatial longitudinal variation (ASLV) clustering. The proposed clustering procedure represents a unique platooning framework since vehicles are not forced into platoons but self-identified as platooned depending on their inter-vehicular longitudinal spatial variation. The proposed concept explicitly considers micro-level vehicle-following behavior and determines inherent discontinuities is traffic patterns, which differentiate platooned over the non-platooned arrivals.

As for the proposed phase scheduling logic, the motivation aligns with the prevailing state of the practice: solve a phase allocation problem in real-time, based on the most reliable high-resolution data obtainable, while assuring applicability in a variety of traffic conditions and

intersection/signal configurations. The method stays away from complex mathematical and statistical models and focuses on operational logic and ease of implementation.

Based on the current signal status and estimated temporal gaps in platoon arrivals, phase allocation is performed such that it minimizes total intersection delay accounting for the marginal cost of serving a specific platoon vs. any other. Control scheme proposed in this study assumes a flexible phasing sequence, which is vehicle trajectories (data)-driven. There is no notion of cycle length as such, the controller is unrestricted, sequence-wise, when scheduling platoon-based phases, provided vehicle-based objective function computations warrant it. Accordingly, phase rotation, skipping, extension and truncation are allowed.

To make sure no vehicle, on any approach, waits excessively, maximum allowable waiting time is imposed on the control parameters optimization. The proposed method accounts for two conflicting objectives: minimizing total delay and maximizing throughput.

The overall procedure consists of the following three sequential stages:

1. Platoon Self-Identification Algorithm
   Identifies platoons formed by the vehicles

2. Phase-platoon Allocation Algorithm
   Identifies the order of serving the platoons and associated signal phases

3. Determining Phase Duration in Real-time
   Determines the duration of the signal phases

7.1.1. Platoon Self-Identification Algorithm

Vehicles traveling in temporal and spatial proximity constitute a platoon. Vehicles traveling in the same direction, scattered among different lanes, can be partitioned into different platoons.

On a multilane approach, vehicles are mostly affected by the vehicles with which they share lane and traffic flow characteristics. Rarely and less significantly, vehicles are affected by the vehicles in adjacent lanes. Furthermore, drivers adjust their positions and speeds in response to vehicles in their immediate vicinity, primarily the vehicle ahead. Considering this to be the standard in a car following behavior, and acknowledging no overtaking is expected near the intersection's cross-section, this study designs a clustering technique which considers the positional adjustment drivers make. This study conducts a movement-based analysis, due to prerequisite compatibility between movements and traffic controller phases. All lanes corresponding to a movement can be observed as a single constitutive element of the control system which allows for headways between the vehicles to be quite small. As a result, it can be inferred that most vehicles tend to keep short headways with their leading vehicles (*64*).

Therefore, the proposed procedure considers, primarily, a vehicle's spatial longitudinal variation, contingent upon the clustering method identifies platoons on each approach. Initially, vehicles are grouped based on their distance from the intersection (cross-section), to eliminate vehicles' lateral position variation on multilane approaches. Again, this reasoning is justified since vehicles are separated per movement, and then grouped longitudinally. In this setup, the vehicles in adjacent lanes, conditioned upon their longitudinal proximity, belong to a single platoon. This concept is visualized in Figure 7-1. Each vehicle's front edge is projected onto the center line of the approach.

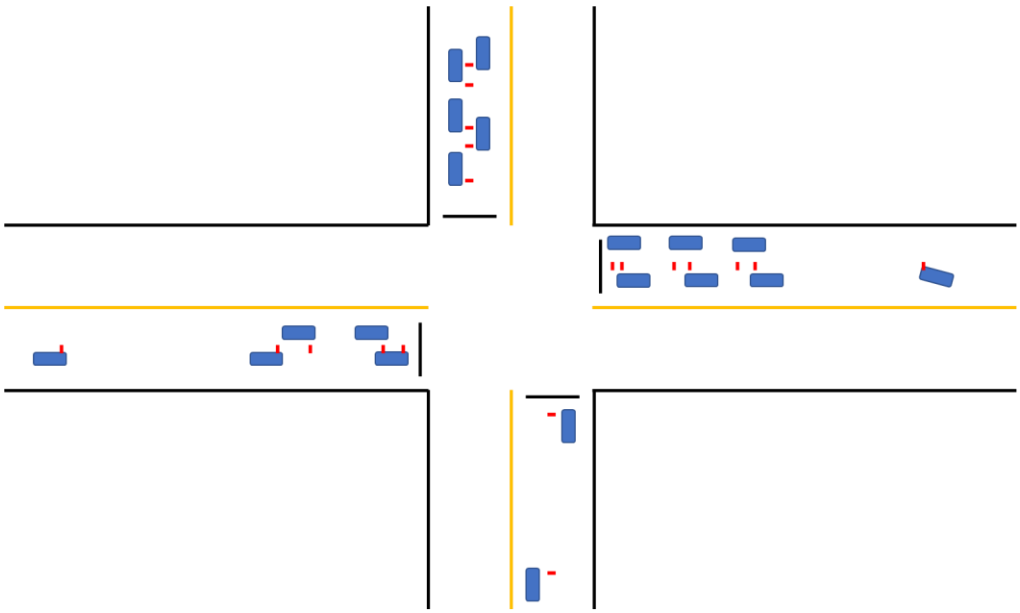


Figure 7-1. Illustration of longitudinal spatial variation

ASLV clustering technique is applied to each movement group of vehicles according to the following steps:

Step 1. Calculate spatial distance from any vehicle (on each approach) to the vehicle closest to the intersection

Step 2. For each vehicle, identify its leading (influencing) vehicle

Step 3. Adjust the position of the following vehicle with respect to its lead vehicle's position

Step 4. Apply the K-means clustering method until convergence criteria are reached

Drivers adjust their behavior according to actions of the vehicle ahead. The preceding vehicle is identified as influencing, if and only if, the distance between the two vehicles (subject and

preceding one) is within the *look ahead distance*. In this manner said the vehicle is verified to exist within reasonable proximity of its preceding vehicle and could have been expected to occupy its position, had there been no lead vehicle.

Look ahead distance is computed as:

$$VehLen + SafeGap * V_f \quad (7\text{-}1)$$

$VehLen$, represents a single (regular) vehicle length to account for the preceding vehicle's length. $SafeGap * V_f$, represents a safety gap between vehicles multiplied by the prevailing speed of the vehicle-follower and accounts for the safe distance between vehicle follower and its lead.

In the presence of influencing vehicle, driver-follower adjusts its position.

Adjustment in the position (*adjustment margin*) is computed as:

$$VehLen + SafeHead * V_f \quad (7\text{-}2)$$

$SafeHead$ represents the safe headway time i.e., desired headway time in seconds between the leader and follower vehicle. In addition to the adjustment margin, the correction includes the lead vehicle's already applied the positional correction. For each follower, this means modifying its current position by subtracting the *adjustment margin* and the correction already applied to the influencing vehicle.

As a result, approach/movement traffic flows are partitioned through ***Adjusted Spatial Longitudinal Variation (ASLV)*** clustering. ***ASLV*** clustering is a type of K-Means clustering repeatedly performed until the convergence criteria are satisfied concerning vehicles adjusted longitudinal positions.

The K-Means clustering method minimizes a within-cluster sum of squares (WCSS). Hence, the observations are divided into **k** clusters such that each observation belongs to a cluster with the nearest mean. In this analysis, **k** was ***NOT*** pre-defined by the user and varies with real-time recorded vehicle positions. Let, **X** be a set of observations ($\mathbf{x_1}$, $\mathbf{x_2}$,…, $\mathbf{x_s}$), where each observation represents an n-dimensional real vector, k-means clustering aims to partition $\boldsymbol{s}$ observations into $\boldsymbol{k}$ ($\leq \boldsymbol{s}$) clusters $C = \{C_1, C_2, \dots, \_C_k\}$ so as to minimize the WCSS given by equation (7-3):

$$\arg\min_{c} \sum_{i=1}^{k} \sum_{x \in C_i} \|x - \mu_i\|^2 \qquad (7\text{-}3)$$

where, μ_i is the mean of cluster $\boldsymbol{C_i}$. The method establishes three criteria of convergence to confirm vehicles have been partitioned into representative clusters.

1. There are fewer clusters than the number of vehicles
2. WCSS is shorter than the average vehicle length
3. Marginal reduction in the WCSS is less than 20%.

The results of the applied clustering method are presented in Figure 7-2. Vehicles are, as mentioned, separated per movement, corresponding to a NEMA standard dual-ring barrier phase configuration (*150*) and clustered accordingly. Each of the platoons is classified by its corresponding approach, phase (movement), platoon ID (platoons are ordered with respect to their arrival times), arrival time, associated total delay, estimated intersection clearing time[1] and end time (exit time when last vehicle has cleared the intersection) as shown in Figure 7-2.

[1] Please refer to Appendix for details.

	Approach	Phase	PlatoonID	ArrivalTime	Delay	ClearingIn	NoVeh	EndTime
1	1	6	1	4.216113	60.63903	17.2657396	17	21.481853
2	1	6	2	11.609657	167.18311	0.0000000	1	11.609657
3	1	1	1	2.511504	140.08346	3.3151746	2	5.826679
4	2	2	1	4.775335	145.66318	14.1613064	14	18.936642
5	2	2	2	4.955721	167.55183	4.7988198	2	9.754541
6	2	5	1	3.440977	108.91293	4.3680229	3	7.809000
7	3	8	1	6.744979	21.07677	0.0000000	1	6.744979

.

.

.

Figure 7-2. Sample result of clustering

7.1.2. Phase-platoon Allocation Algorithm

Identified platoons are served on a quasi-first come first serve basis. For each of the identified platoons, corresponding arrival times are determined. Arrival time is referred to as the time of arrival of the platoon's head vehicle at the approach stop bar. For example, queued vehicles (at the stop bar) will constitute a platoon whose arrival time is 0.

Platoons arriving within 5 seconds of the earliest arriving platoon are chosen as candidate platoons to be served. This subset of platoons forms the set $\boldsymbol{CanP}$.

Given:

$\boldsymbol{CanP}$: Set of platoons approaching the intersection

$T_i^A, \forall\, i \in \boldsymbol{CanP}$: Arrival times of the identified platoons

$T_i^E, \forall\, i \in \boldsymbol{CanP}$: Exit times of the identified platoons

Parameters:

$\boldsymbol{m}_\phi$: performance index of platoon ϕ

Objective:

$$\phi^* = \underset{\phi \in CanP}{\operatorname{argmin}} \ \boldsymbol{m}_\phi \qquad (7\text{-}4)$$

Subject to:

$$MC_\phi^\gamma = \begin{cases} \left(T_\phi^E - T_\gamma^A\right) N_\gamma & if\ T_\phi^A \geq T_\gamma^A \\ 0 & otherwise \end{cases} \qquad (7\text{-}5)$$

$$MC_\phi = \sum_{\gamma \in \boldsymbol{CanP} \setminus \phi} MC_\phi^\gamma \qquad (7\text{-}6)$$

$$\boldsymbol{m}_\phi = \left(\frac{MC_\phi + \sum_{\gamma \in CanP \setminus \phi} Delay_\gamma}{Delay_\phi} \right) \qquad (7\text{-}7)$$

Vehicles respect the laws of physics

Drivers follow the Wiedemann 74 car following model

No frequent transitions between signal plans so as not to disrupt the progression

Vehicles on the unserved vehicles are not waiting for more than $\boldsymbol{T_{Max}}$

Considering the marginal cost of activating a phase in addition to the total delay experienced by the vehicles, the next phase is determined. $\boldsymbol{m}_{\phi}$ represents system's performance if platoon ϕ is to be served. Procedure to determine $\boldsymbol{m}_{\phi}$ is described below. $Delay_v$ is the delay experienced by the vehicle v. For CAVs it is the accumulated delay along the corridor and for regular vehicles, the delay experienced at the intersection approach under consideration.

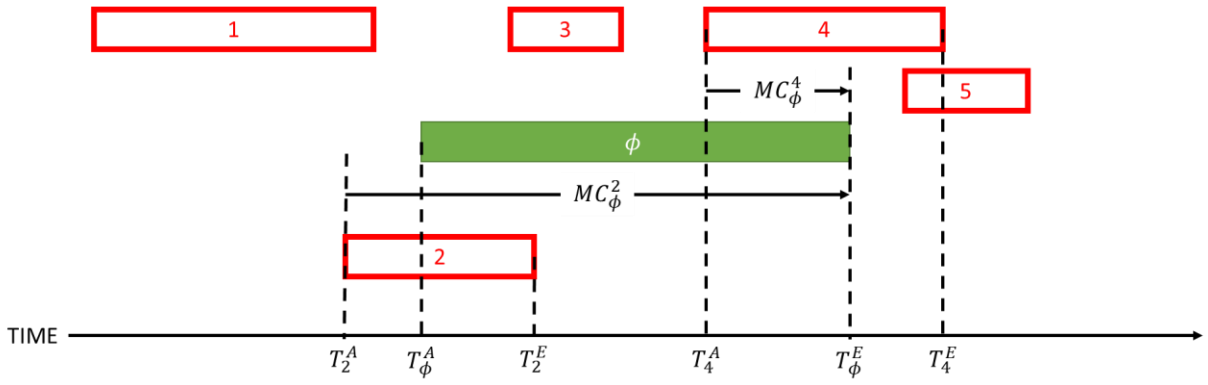


Figure 7-3. Visualizing marginal cost.

Marginal cost, MC_{ϕ}^{γ}, is defined as the additional delay incurred by a non-served platoon γ due to the right of way afforded to platoon ϕ. Computation of marginal cost depends on the platoon service/clearing time. Service time is defined as the time duration between the arrival of the first vehicle and exit of the last vehicle in the platoon. Detailed calculations are provided in the APPENDIX I. T_p^A and T_p^E are the arrival and exit time of a platoon p. N_p is the number of vehicles in platoon p.

Referring to the example presented in Figure 7-3, marginal cost is computed per case:

Case 1: An unserved platoon is expected to arrive or exit earlier than the one served. Therefore, an unserved platoon must wait for the entire served platoon phase duration (MC_{ϕ}^{1}, MC_{ϕ}^{2}, MC_{ϕ}^{3}).

Case 2: An unserved platoon is expected to arrive or exit later than the considered platoon. Consequently, it will only wait for the overlapping time duration (MC_{ϕ}^{4}).

Case 3: An unserved platoon is expected to arrive later than the considered platoon. Hence, it incurs no marginal cost (MC_{ϕ}^{5}).

Marginal cost aggregated over all unserved platoons gives the total marginal cost of serving the platoon ϕ, MC_{ϕ}, and is defined in equations (7-6) and (7-7).

Marginal cost together with the total platoon delay governs the "optimal" next phase selection. To determine which phase-platoon pair to service next, out of the $\boldsymbol{CanP}$ set of alternatives, a tradeoff is calculated between the incurred delay and time savings if one platoon is to be discharged instead of another. Total delay $Delay_p$ of platoon p is defined as the sum of individual vehicle delays. CAVs are assumed to log the delay along their route and this accumulated delay is considered towards total platoon delay calculations. The platoon, ϕ^*, for which the ***objective function***, presented in *equation* (7-4) the value is minimized will be served next.

The phase associated with platoon ϕ^* is referred to as ζ'. Considering NEMA-RBC non-conflicting phasing combinations (*150*), the algorithm then continues to find a potential concurrent phase, ζ'', if warranted by the demand. If multiple phases have the same associated objective

function value, the one serving the largest number of vehicles will be preferred. Referring to Figure 7-3, platoon 3 can be served along with ϕ platoon as long as their phases are not conflicting. To determine if multiple platoons that can be served with ϕ^*, the optimization problem is again solved after updating the marginal cost and objection function values as per equation (7-8) and (7-9). Accordingly, ϕ^{**} is identified as the platoon to be served along with ϕ^*. The phase associated with platoon ϕ^{**} is referred to as ζ''.

$$\boldsymbol{m}'_{\phi} = \left(\frac{MC_{\phi} + \sum_{\gamma \in CanP \setminus \{\phi, \phi^*\}} Delay_{\gamma} - MC^{\phi}_{\phi^*}}{Delay_{\phi} + Delay_{\phi^*}} \right) \quad (7\text{-}8)$$

$$\phi^{**} = \underset{\phi \in CanP \setminus \phi^*}{\operatorname{argmin}} \boldsymbol{m}'_{\phi} \quad (7\text{-}9)$$

Once ζ' and ζ'' are identified, the actual phase duration is determined based on vehicle trajectories in real-time, monitored every second.

7.1.3. Determining Phase Duration in Real-time

Once the phases are identified, the algorithm (in real-time) continuously checks, with the frequency of one second, whether enough phase time has been allotted to a platoon being served, given that additional vehicles might arrive before the phase has terminated. If this is the case, the green will be extended to service the incoming vehicles, within a prespecified distance from the intersection. Conversely, if the platoon has cleared the intersection sooner than expected, the green terminates.

These control parameters are defined under restrictions such as upper bound on green time and distance-based extension time. Distance-based maximum green is established to limit the amount of time allocated to a phase, to avoid excessive waiting times at unserved approaches/movements. Also, a distance-based gap out (adjustment margin) defines whether the newly arriving vehicles will be served during the current green or stopped to wait for the next green indication.

Four basic variables characterize each platoon: size (number of vehicles in the platoon), platoon intra and inter-headway or location (of the head and tail vehicle) and speed (of individual vehicles in the platoon), from which the arrival and intersection clearing times can be estimated. Please note, that intersection clearing time was computed based on standardized values of signalized approach discharge headway (if stopped to account for start-up lost time), current speed and allowed acceleration rates to estimate corresponding phase's duration. Platoon service time is determined as the time elapsed between the arrival of the first vehicle and the exit time of the last vehicle in the platoon. This duration gives an estimate of the actual phase duration.

7.2. Numerical Experiments Results

The control strategy developed in the study is a platoon-based heuristic that aims not to delay and interrupt platoons' progression. It is considered preferable to clear the vehicles before the platoon arrives at an intersection but whether this is possible depends on the conflicting movements' vehicle and signal state. Demonstrating the effectiveness of such a control concept was one of the primary goals of this research.

Presently, the lack of available connected vehicle information is an issue when relying on such data signal control-related applications. For this reason, conceptual framework, experiments, and evaluation in a mixed vehicle environment of automated, connected and regular vehicles, with high-resolution vehicular information, were conducted in a microscopic simulation setting. A real-world intersection, of SR7 and Broward Blvd, in Fort Lauderdale, Florida, was modeled and calibrated in VISSIM (*148*), to represent field conditions as realistically as possible. Actual time of day (TOD) signal timing plans were utilized as the baseline state of the traffic signal system. Then, the robustness and effectiveness of the proposed control method were examined under different mixes of traffic and demand levels.

Demand levels were divided into three categories: low, medium and high, each being 30 minutes long. A 2-hour long simulation horizon was created to represent the demand build up from low (off-peak conditions), then medium to oversaturated (morning or evening peak) and then reverting to medium, to characterize recovery after oversaturation dissipates. Directions East and West (Broward Blvd.) are major approaches with 100% demand factor of the corresponding demand level. Please refer to the representation of traffic mixes and demand levels in Table 7-1 and Table 7-2, respectively. Overall, four different ***C***onnected, ***A***utomated and ***R***egular (***CAR***) vehicle mixes were tested over a 2-hour long horizon with three demand levels. Additionally, a segment of 6 intersections was also modeled to as close to the reality as possible emulating the field conditions.

Table 7-1 Traffic composition.

Case	Traffic mix		
	C	**A**	**R**
1	0.0%	0.0%	100.0%
2	0.0%	33.3%	66.7%
3	33.3%	100.0%	66.7%
4	0.0%	66.7%	33.3%
5	33.3%	33.3%	33.3%
6	66.7%	0.0%	33.3%
7	0.0%	100.0%	0.0%
8	33.3%	66.7%	0.0%
9	66.7%	33.3%	0.0%
10	100.0%	0.0%	0.0%

Table 7-2 Traffic demand levels.

Demand Level	
Case	**Flow**
Low	1800
Medium	2400
High	3600

Demand Factors	
Dir	**%**
EB	100%
WB	100%
NB	45%
SB	45%

The control strategy was tested during consecutive demand periods for each ***CAR*** mix and assessed against the corresponding baseline of SYNCHRO (*152*) optimized TOD signal timing plans. The entire 2-hour testing horizon was divided into eight 15-minute demand periods. This is to say that demand periods 1 and 2 pertain to the low demand level, the next two correspond to the medium demand level and so forth.

The platoon-phase allocation algorithm was applied to two platooning techniques – ASLV and simple platooning, to examine and isolate the effect of ASLV platooning. Simple platooning represents a conventional gap-out platooning method, which utilizes a critical headway threshold of 2.5 seconds to distinguish between platoons of vehicles. Vehicle-actuated control logic was the baseline in this study. The reported results in the figures below, present the three cases comparatively. With respect to the baseline (black numbers in the center of the figure), operational success of the other two strategies is presented in relative terms – percent change in the corresponding performance metric relative to the ground truth value. Green (orange) bars mark the improvement of the ASLV (simple) platooning strategy compared to the base case. Similarly, brown (red) colored portions in the charts indicate the worsening in performance indicators, respectively.

The results are reported at two levels – isolated intersection level and arterial corridor level. Overall, at both levels – a similar trend is observed referring to each measure of effectiveness. The two platooning methods seem to be performing similarly, with the ASLV method offering incremental improvements over the simple platooning strategy.

As mentioned earlier, the signalized corridor testbed deploys the ASVL platooning and advanced control logic at only one intersection. Isolated intersection results offer more significant improvements in performance, compared to the corridor-level, which is to be expected. Aggregation of results, at corridor level, cancels out some of the positive effects achieved at an individual intersection level.

Figure 7-4 through Figure 7-9 illustrate the state of the system after the strategy is applied, against specific case's ground truth, i.e., baseline conditions and the conventional gap out platooning

strategy. Several MOEs were considered to evaluate the performance of the ASLV platoon-based control strategy.

For the isolated intersection case, the entire 2-hour testing horizon was divided into eight 15-minute demand periods. For the whole corridor, the testing horizon was divided into 10-minute intervals.

Referring to Figure 7-4, the positive effect of the strategy on queue length reduction is evident. Except CAR mix Case 1 (100% regular vehicles) during the second oversaturated period (demand period 6), when increment in queue length is observed, in all other cases, the strategy significantly outperformed the no-strategy alternative. The strategy proves to be effective in reducing the delay as well, as can be observed in Figure 7-5. Total delay reduction is found to be significant in the isolated intersection case, especially during oversaturation, for both strategies tested. In some cases, even higher with the traditional platooning framework applied – as high as 50% reduction. However, the reduction in total delay achieved at the corridor level was up to 7% (8% with traditional platooning applied). To isolate the effect of congestion delay from stopped delay, the effect of the strategy on stopped delay was also studied. The trend in the stopped delay was similar to that of total delay, but with a higher magnitude. The strategy allowed for a reduction in stop delay of up to 8% at the corridor level. More significant improvements were achieved during oversaturated traffic conditions. The improvement in the quality of service was realized even in highly congested conditions.

At corridor level, conventional gap-out platoon-based control, unlike the ASLV self-identification control method, fails to consistently achieve superior operational efficiency compared to the vehicle-actuated type of control (e.g., Figure 7-7 through Figure 7-9).

At an isolated intersection level - period 6 – the second oversaturated period – proves to be the most challenging, as expected. Both strategies underperform compared to the vehicle actuated base-case, for traffic mixes 1 and 6. While it is reasonable to assume the strategy might underperform when the traffic stream consists of regular vehicles only (Case 1), it is counterintuitive that the same occurs when 66 % of the fleet is connected and 33 % regular vehicles. Such an outcome could be attributed to the presence of regular vehicles which reduces the estimation accuracy of platoon arrival/exit times.

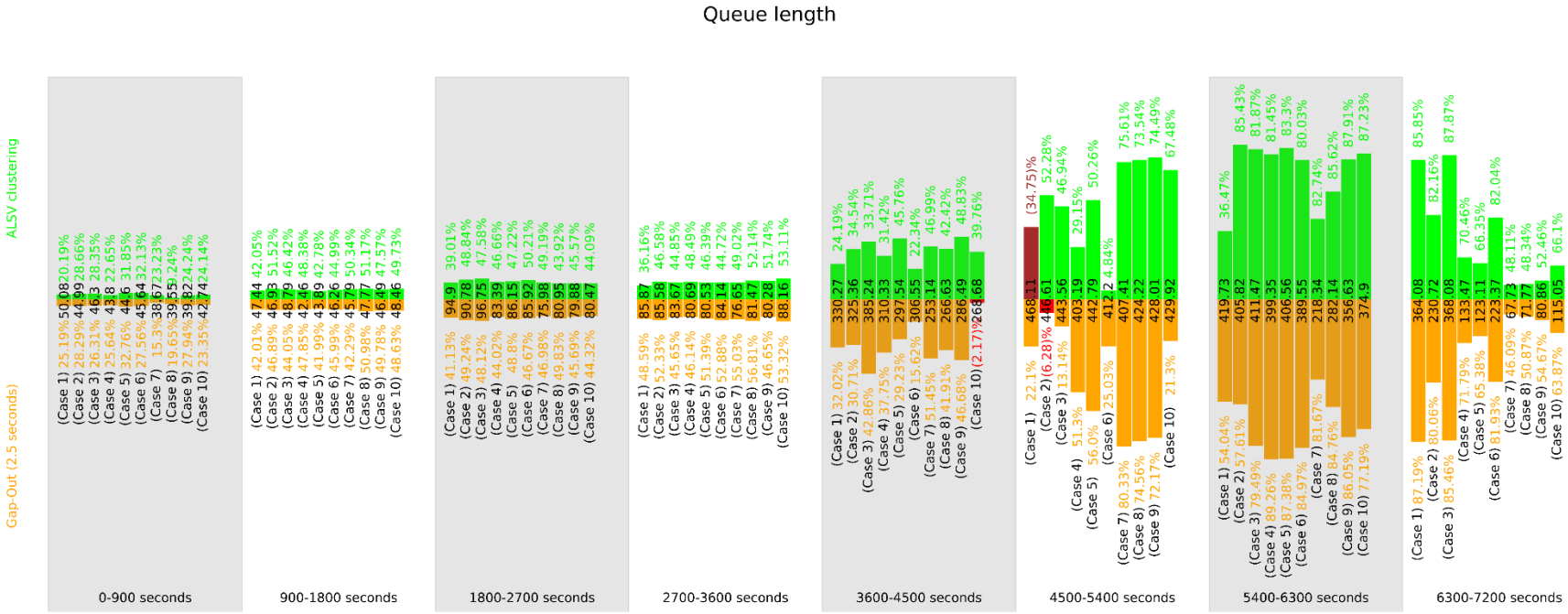


Figure 7-4. Impact of control strategy on queue length – isolated intersection.

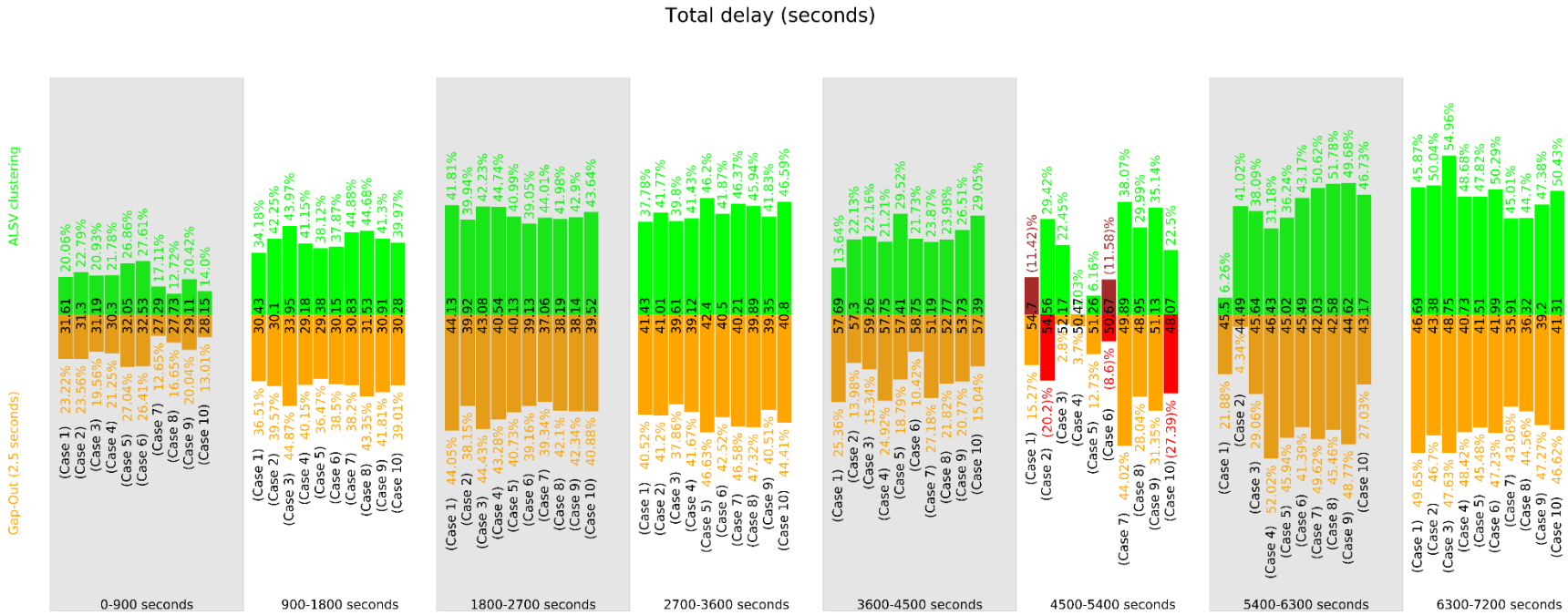


Figure 7-5. Impact of control strategy on total delay – isolated intersection

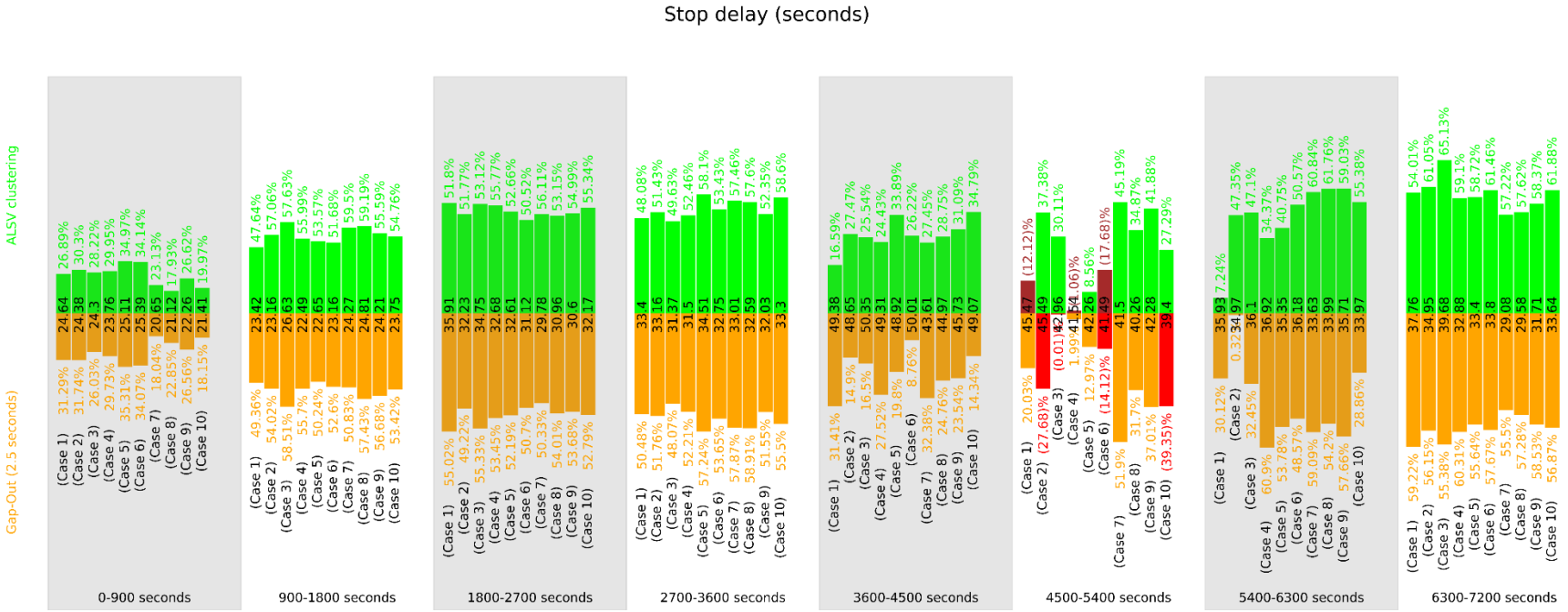


Figure 7-6. Impact of control strategy on stopped delay – isolated intersection

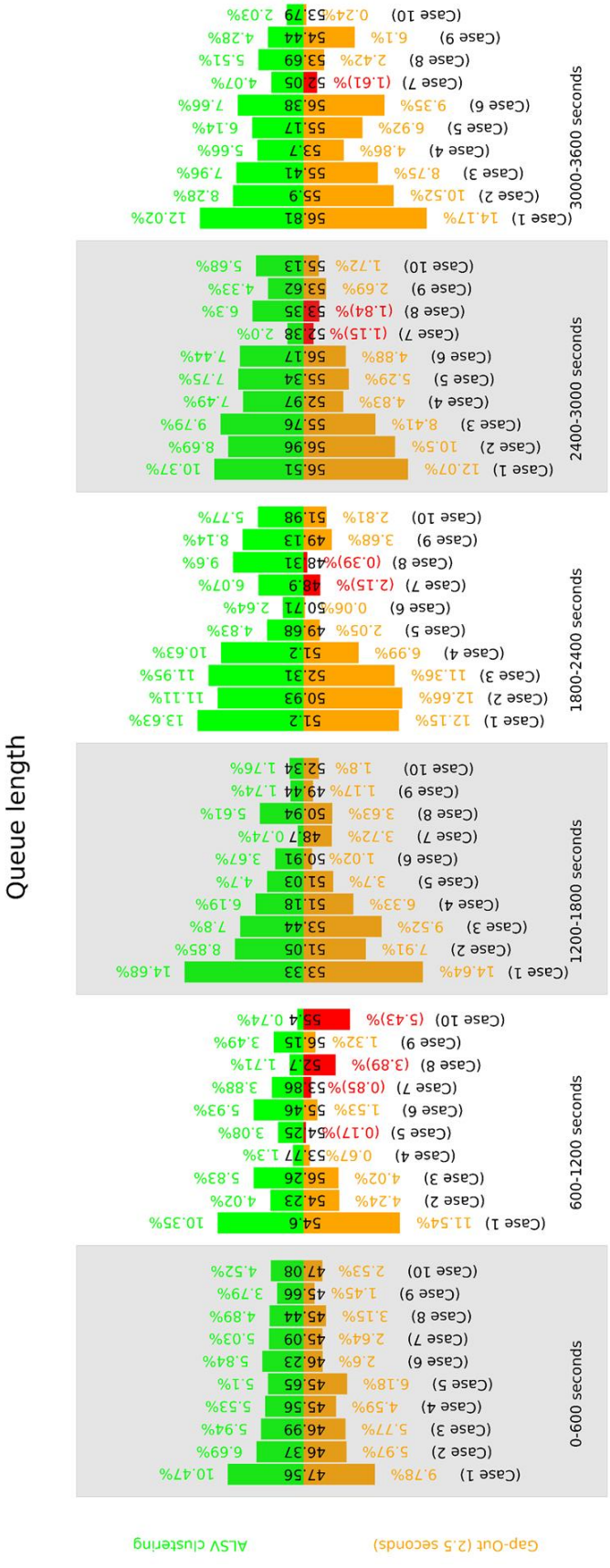


Figure 7-7. Impact of control strategy on queue length – corridor

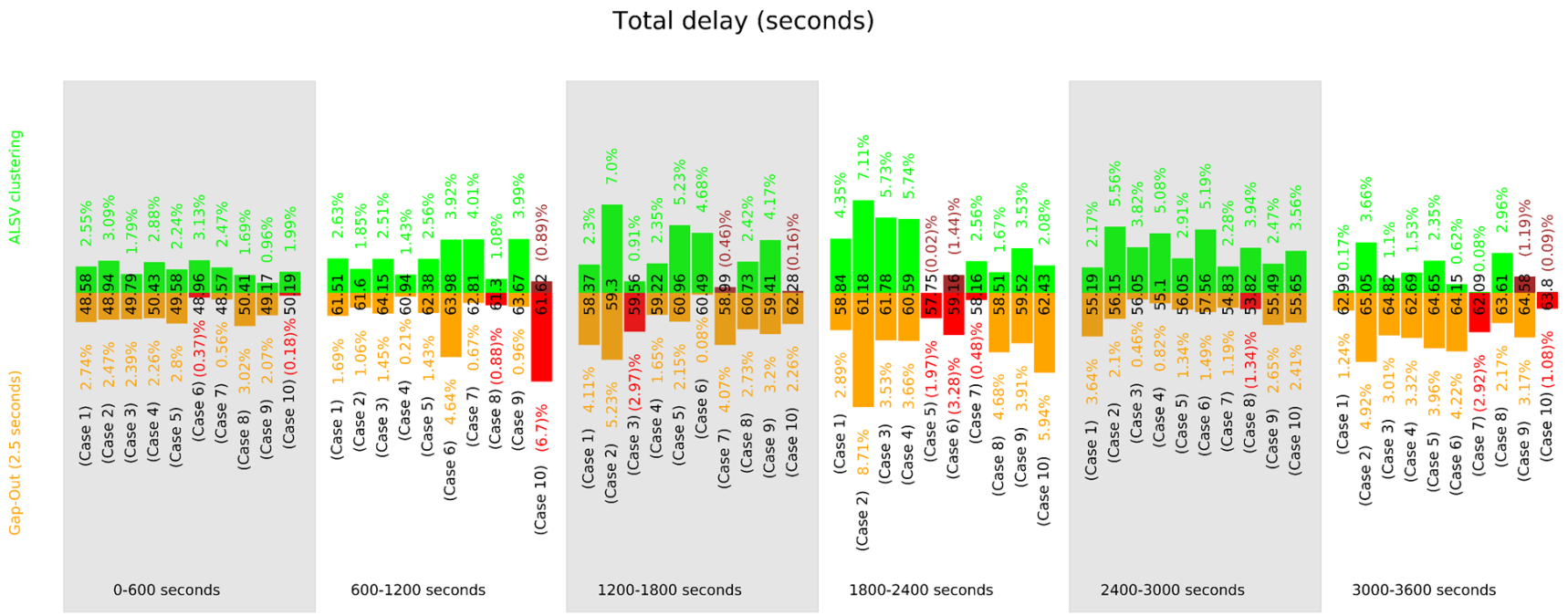


Figure 7-8 Impact of control strategy on total delay – corridor

Figure 7-9 Impact of control strategy on stopped delay – corridor

7.3. Conclusions

This section discussed the design of an efficient generic adaptive signal control algorithm for mixed vehicular traffic environments aimed at improving the quality of service, given the objectives - minimizing delay and maximizing throughput. To this end, an intelligent fully-actuated controller logic computes vehicle-based performance metrics to optimize control parameters in real-time. CAVs are again, as in Chapter 6, assumed to transmit their accumulated delay thus enabling vehicle progression when warranted by the incurred delay. The real-time signal control scheme presented incorporates the concept of marginal delay cost in its objective function calculation, i.e., accounts for the unserved vehicles' extra delay for any additional vehicle served.

Most of prediction-based signal control strategies fail due to the discrepancy between the predicted/estimated and real-world traffic conditions. Obtaining relevant information at the right time should ensure efficiency of the adaptive control schemes (*153*). The proposed strategy addresses this issue successfully. Most of the strategies found in the literature revolve around merely deciding whether to switch or extend the current green to accommodate the incoming platoon. Conversely, the proposed method puts forward a control logic which proactively determines the next optimal phase-platoon to serve and then continually adjusts its phase duration in real-time reacting to the prevailing flow patterns.

Its distinctive features are:

- Platoon self-identification technique, which delineates platoons based on vehicles adjusted spatial longitudinal variation.

- Phase-platoon sequencing which accounts for total delay experienced by vehicles on each approach and the marginal cost of affording the right of way to a platoon vs. another. Please note that total delay includes CAVs accumulated delay along a route
- Phase duration is effectively adjusted in real-time based on actual vehicle arrivals on a second-by-second basis.

The focus of this research effort was to develop a control methodology that will provide progression on arterial roads through improved local intersection control, i.e., platooning. Operational efficiency at a corridor level was achieved by identifying natural breaks in traffic flow and consequently, grouping vehicles into platoons. The concept of platooning in this research deviates from what the traditional transportation research establishes. A notion of adjusted spatial longitudinal positioning is devised to scrutinize for the standard in driving behavior, i.e., that a driver will innately correct its position in reaction to its leading vehicle's actions. If there is no gap in the incoming traffic, such vehicles should be considered as a single entity – a platoon. Adjustment margin was established to account for the position correction of vehicles on an approach, had there been no lead/preceding vehicle. The assumption is that the vehicle would have positioned itself in its preceding vehicle's place on the link/lane. The optimal number of clusters, i.e., platoons (per approach) is demand-driven, therefore, independent of any (critical) threshold such as critical inter-vehicle spatial or temporal headway, cumulative headway, (minimal) number of vehicles constituting a platoon, etc.

Benefits cited in most previous studies' do not relate to the real-world representative geometric, traffic and control characteristics. In reality, these configurations can be complex, which poses a challenge to the adaptability of the proposed methods. Therefore, the proposed strategy was tested

on a real-world signalized intersection, albeit using microsimulation. The adaptive control method developed, although unique in its design logic, information utilized, objective function and modeling approach apply the fundamental NEMA-RBC controller standard of non-conflicting movements grouping. (*150*)

CHAPTER 8. PATH FUNDAMENTAL DIAGRAMS

As it was already discussed, the type of facility influences the driving behavior and vice versa. As demonstrated in Chapter 4, Chapter 6, and Chapter 7, the impact of connectivity, as well as the effectiveness of strategies, was examined, separately, and conclusions drawn were facility-based. Considering the dynamic nature of demand patterns and hard to predict driver's behavior, identify and quantify the differences in driving behavior and traffic stream variables' relationships at a higher level was critical. Fundamental diagrams are the basis of such studies. Fundamental diagrams are known to hold at the link level and a network level. The next question that arises is that do fundamental diagrams hold at the path level. Paths are defined by the drivers. The effort in this chapter is aimed at answering this question. This chapter also offers a new perspective on transportation system partitioning based on user-defined paths' flow relationships.

Advances in the theoretical sophistication of available urban transportation models have steadily increased over the decades. Modeling approaches progressed considerably since the original, static and mainly, aggregate *four-step model*s dating between the1950s and 1960s; to the *disaggregate demand* and *network equilibrium* extensions in the 1970s and 1980s; and then the recent advancements with *dynamic simulation* models of the 1990s and 2000s. Along with theoretical advances, practices were also improved and deployed in the field to achieve congestion relief. A few successful attempts include a regional scale pricing policy deployed in London and Stockholm, based on a pre-determined time of day pricing scheme. Recently, state-dependent pricing has been researched and successfully tested in simulated environments for real-time freeway management. Concerning arterial corridors, traffic signals have also been reasonably successful in providing efficient vehicle progression. Traffic control systems were upgraded

considerably since fixed types of control - UTCS (Urban Traffic Control System, FHWA, 1970s) Through traffic responsive to real-time adaptive control such as SCATS (*154*), SCOOT (*155*), OPAC (*156*), UTOPIA (*157*), RHODES (*158*), ACS-Lite (*159*), MOTION (*160*), and others.

With technological and conceptual advances came more sophisticated simulation models. These models have been successfully implemented to simulate and predict traffic flow characteristics of multi-modal transportation networks. However, field practices based on control strategies and state prediction models are still far behind (*20*). Congestion is a dynamic, large-scale and heterogeneous problem. Furthermore, non-recurrent traffic events like accidents or other unprompted events are, in their nature, unpredictable. The consequences of which might be chaotic oversaturation. Hence, there is a need for a practical study that can address dynamic demand patterns and hard to predict driver's behavior. It is essential to isolate the problem and investigate it at the right level.

The fundamentals of the proposed framework consist of understanding and characterizing drivers' behavior at a microscopic level, modeling traffic flow at a mesoscopic level, then identifying popular user-defined paths, i.e., between O-D pairs, and clustering them based on their traffic flow relationships. This approach, as already mentioned, reduces the computational burden while maintaining as much network detail as possible.

The traffic network, for this research, was subdivided based on the route choice behavior, which represents a novel approach to network path clustering. Also, a new clustering technique was designed and demonstrated to outperform the existing methods.

8.1. Data Description

Simulated vehicle trajectories were used to conduct the analysis. Trajectories were obtained from the Chicago City network simulation model, where operational conditions were established as the clear weather with no precipitation and normal visibility.

Chicago testbed was extracted from the entire Chicago Metropolitan Area network to enhance the estimation and prediction performance during the implementation procedure. This testbed represents a section of the network previously described in Chapter 4 that was utilized to evaluate the impact connected vehicles had on operational efficiency at a network level.

Obtained vehicle trajectories comprised of the entry and exit time of vehicles for each link traversed. At each link, $\boldsymbol{link}$, a vehicle, $\boldsymbol{veh}$, was assumed to be travelling with a constant speed, V_{veh}^{link}, that was determined by the length of the link, L_{link}, divided by the time a vehicle spent on that link, T_{veh}^{link}. To eliminate the errors due to (at origin nodes) demand sources and similarly (at destination nodes) demand sinks, the first and last link were not considered in the analysis. Mathematical formulation is presented below:

$$V_{veh}^{link} = \frac{L_{link}}{T_{veh}^{link}} \qquad (8\text{-}1)$$

As previously stated in the conceptual framework, Chapter 3, link-based measurements are necessary to obtain network-wide relationships among flow, density/occupancy, and speed. Since obtaining vehicle-trajectories will become readily available in the near future, applying the approach designed by Saberi et al. (*1*) seemed warranted. The authors demonstrated that trajectory

based measurements of network-wide relationships provide higher accuracy as compared to that of link-based (*1*).

8.2. Methodology

The methodology adopted in this chapter is divided into three modules (Figure 8-1).

Module 1 represents a mesoscopic DTA tool. DYNASMART-P, a mesoscopic dynamic traffic assignment simulation tool, was used to simulate traffic dynamics on a network level. Further tool specifications and relevant resources are provided in Chapter 4.

Module 2 designs the procedure for identification of the popular paths taken by the drivers. It is based on the vehicle trajectories output from Module 1. With richer data availability and associated information processing technologies, it is possible to track drivers' trip trajectories through probe data from GPS or smartphones. Popular paths are defined as the paths (or sub-paths) shared by a considerable number of travelers given spatial and temporal boundaries. To identify popular paths a data-driven clustering method for trajectories was adopted. The core part takes advantage of the network topology to detect common links in a road network with archived vehicle trajectory data. The detailed problem formulation and algorithmic procedures for trajectory clustering have been described in Hong et al. (*161; 162*).

Module 3 comprises clustering of the Path Fundamental Diagrams (PFDs). PFDs were obtained for each of the identified popular paths. For a single path, flow, density, and speed were computed, as described in section 8.1 and 3.3, for all links that constitute said path and was aggregated over 5-minute intervals.

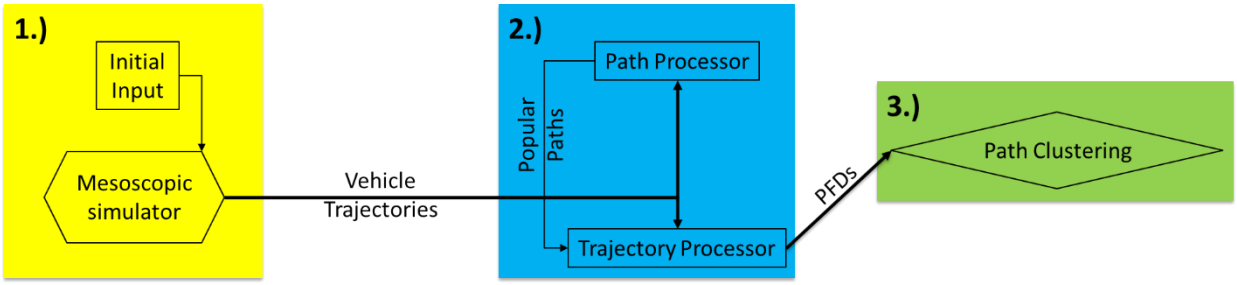


Figure 8-1 Schematics of the methodology adopted in the study.

Clustering PFDs and their characterization is the focus of this chapter's effort (Figure 8-2). The following subchapters, along with the one proposed, describe the clustering methods that were widely exploited according to the relevant literature (and were also examined in this research). The first two methods can be considered conventional and will be discussed in brief. A third method is a novel approach established by the author. The first two methods were implemented in *R language* (*163*) while the third one was implemented in *AMPL* (*164*).

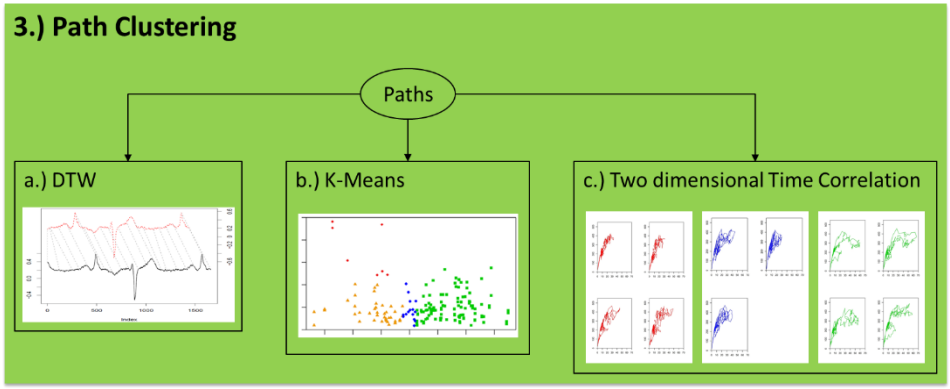


Figure 8-2 Schematics of module 3.

8.2.1. Method 1: Dynamic Time Warping (DTW) Clustering

DTW is used to compute the distance and alignment between time series. Through DTW, the time axis of the series is distorted to measure the similarity between the series. The concept is illustrated in Figure 8-3 below. The gray dashed lines show the mapping between the two time-series.

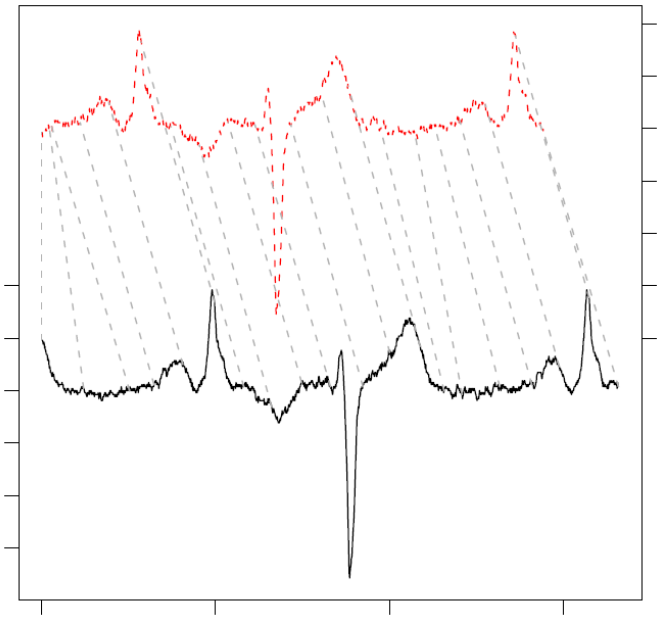


Figure 8-3 Graphical representation of DTW technique. (Source: Giorgino (*165*))

The partitioning was performed on the remapped data, to identify the clusters of the paths. Time series of the velocity of paths were clustered in this method. The mathematical representation of the method is as follows.

Let $\mathbf{x_1}$ and $\mathbf{x_2}$ be the two time-series of length $\boldsymbol{n}$ and $\boldsymbol{m}$ respectively. Please note that, for this study, $\boldsymbol{n}$ and $\boldsymbol{m}$ are equal.

$$DTW_{\phi}(x_1, x_2) = \sum \frac{m_{\phi} lcm(k)}{M_{\phi}} \, , k \in \phi \qquad (8\text{-}2)$$

$$lcm(i, j) = \left(x_{1i} - x_{2j}\right)^2 \qquad (8\text{-}3)$$

Where $lcm(k)$ is the local cost matrix of size $\boldsymbol{n}$ x $\boldsymbol{m}$, and is computed as defined in the equation (8-3). *Warping curve*, ϕ, is the optimal path which is obtained with the following objective.

$$\arg\min_{\phi} DTW_{\phi}(x_1, x_2) \qquad (8\text{-}4)$$

Where $\phi(k) = \left(\phi_{x_1}(k), \phi_{x_2}(k)\right)$ with $\phi_{x_1}(k) \in \{1, \dots, n\}$ and $\phi_{x_2}(k) \in \{1, \dots, m\}$. As defined in the original work, m_{ϕ} is a per-step weighting coefficient and M_{ϕ} is the corresponding normalization constant (*165*).

This analysis was performed in *R language* using the "***tsclust***" command which is a part of the "***dtwclust***" library. Readers are strongly encouraged to read the documentation provided with the library and references therein. Another crucial reading reference about the technique is the work by Giorgino (*165*).

8.2.2. Method 2: K-Means (KM) Clustering

The KM clustering method minimizes a within-cluster sum of squares (WCSS). Hence, the observations are divided into **k** clusters such that each observation belongs to the cluster with the nearest mean. In this analysis, **k** was pre-defined by the user. Let, **X** be a set of observations ($\mathbf{x_1}$,

$\mathbf{x_2}, \ldots, \mathbf{x_s}$), where each observation represents a n-dimensional real vector, k-means clustering aims to partition the $\boldsymbol{s}$ observations into $\boldsymbol{k}$ ($\leq \boldsymbol{s}$) clusters $C = \{C_1, C_2, \ldots, C_k\}$ so as to minimize the WCSS (8-5).

$$\arg\min_{c} \sum_{i=1}^{k} \sum_{x \in C_i} \|x - \mu_i\|^2 \qquad (8\text{-}5)$$

μ_i is the mean for $\boldsymbol{C_i}$. This analysis was performed in *R language* using the "***kmeans***" command which is a part of the "***stats***" library.

8.2.3. Method 3: Two-Dimensional Time Correlation (TDTC) Clustering

TDTC is a novel clustering method that was formulated to cluster the paths based on the temporal correlation of flow and density. Parameters, variables, problem statement, and the constraints are presented and discussed below.

Parameters:

c: Clusters $\in [1,2,3, \ldots, centers]$

p: Paths $\in [1,2,3, \ldots, paths]$

t: Set of time intervals $\in [1,2,3, \ldots, intervals]$

$Data_p$: PFD data

$Dist_p$: Distance of path p to its corresponding center matrix

Variables:

δ_{cp}: Probability that path p belongs to cluster c

Z_p: $\begin{cases} 1 & , if\ path\ p\ is\ a\ center \\ 0 & , otherwise \end{cases}$

Minimize:

$$\sum_{p=1}^{paths} Dist_p \qquad (8\text{-}6)$$

Subject to:

$$Dist_p = \sum_{c=1}^{paths} \left[\left[\sum_{t=1}^{intervals} \left\| [Data_c]_t - \left[Data_p\right]_t \right\| \right] \delta_{cp} \right]; \forall\ p \in [1,2,3,\dots,paths] \qquad (8\text{-}7)$$

$$\sum_{c=1}^{centers} \delta_{cp} = 1; \qquad \forall\ p \in [1,2,3,\dots,paths] \qquad (8\text{-}8)$$

$$\sum_{c=1}^{paths} Z_c \geq centers \qquad (8\text{-}9)$$

$$0 \leq \delta_{cp} \leq Z_c; \qquad \forall\ p \in [1,2,3,\dots,paths]\ , c \in [1,2,3,\dots,paths] \qquad (8\text{-}10)$$

The constraint presented in the equation (8-7) is the probabilistic Euclidean distance of path's distance from a cluster's center. $\|\cdot\|$ represents the Euclidean distance which is then multiplied by δ_{cp}, i.e., the probability that a path, $\boldsymbol{p}$, belongs to a cluster, $\boldsymbol{c}$. Equation (8-8) constraints the sum of probabilities (of a path) across all clusters to add up to one. Please note that, $\boldsymbol{Dist_p}$ and $\boldsymbol{Center_c}$ are two-dimensional matrices. $[\boldsymbol{Center_c}]_{\boldsymbol{t}}$ and $\left[\boldsymbol{data_p}\right]_{\boldsymbol{t}}$ contain density and flow measurements for time step $\boldsymbol{t}$ and center $\boldsymbol{c}$ or path $\boldsymbol{p}$, respectively. $\boldsymbol{Center_c}$ is calculated as the average of data from all member paths of the clusters. This analysis was performed in *AMPL* using *KNITRO* solver due to the existence of nonlinear constraints.

A simulation-based analysis was conducted to test the applicability and effectiveness of the framework developed. The following section first presents the outcomes of each of the clustering techniques tested, individually. It was observed that each clustering technique produces four representative clusters, which may result in the similarity of clustering methods applied and the idiosyncratic features of the data itself. Accordingly, a comparison study was conducted to delineate the differences between the three series of results. Furthermore, common findings from the study are presented towards the end of the section.

8.3. Numerical Experiments Results

8.3.1. DTW clustering

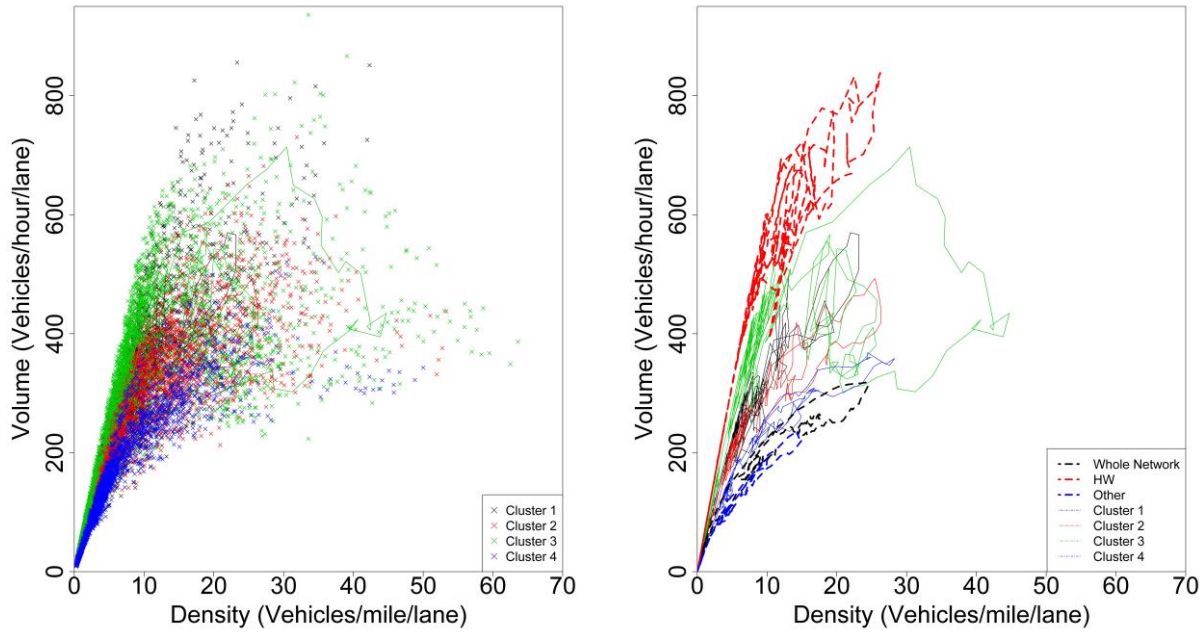


Figure 8-4 DTW based a) NFDs for all the paths with four centers; b) centers of 4 clusters with NFDs of different road types.

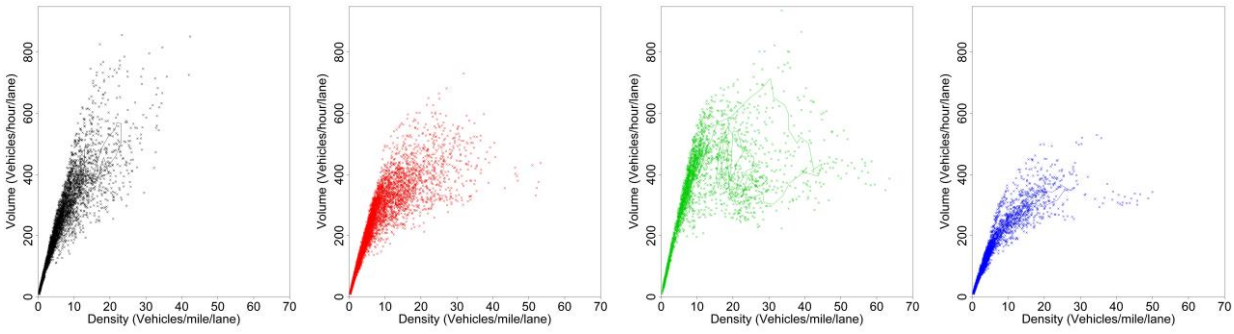


Figure 8-5 NFDs for all the paths by DTW based clusters.

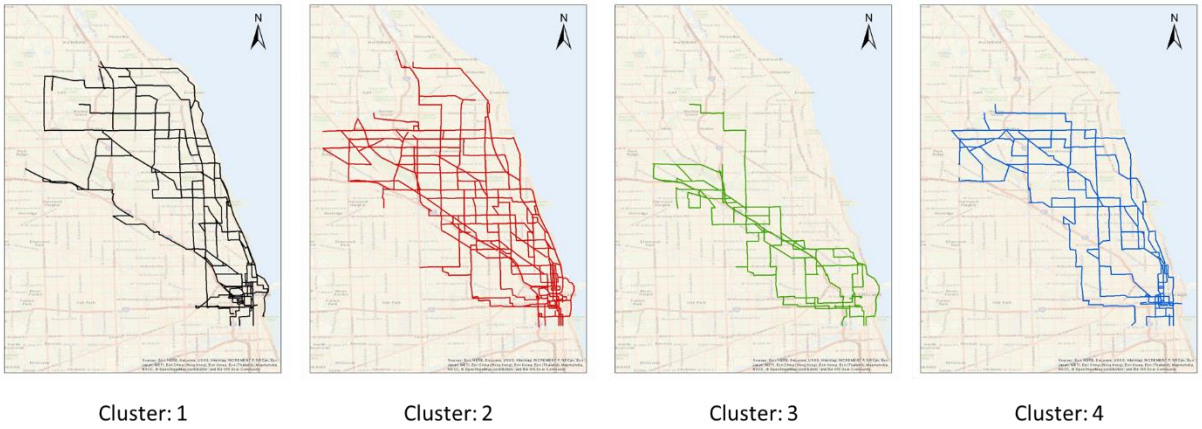


Figure 8-6 Spatial distribution of the paths by DTW based clusters.

Results based on DTW clustering are presented in Figure 8-4 through Figure 8-6. As discussed previously, time series of velocities were used to identify the path clusters. Four significantly different clusters were obtained whose characteristics are as follows:

- *Cluster 1*: Slightly congested, high flow, (representing primarily) arterial roads with segments of Lakeshore Drive highway
- *Cluster 2*: Moderately congested, a mix of arterial and FWs
- *Cluster 3*: Highly congested, but maintains the flow rate, big hysteresis loop, mainly FWs, in this case, hysteresis loop vertical distance is the largest
- *Cluster 4*: Moderately congested, low flow rate when congested, primarily the HWs

It is interesting to note that a spatial pattern emerges from the clusters.

8.3.2. KM clustering

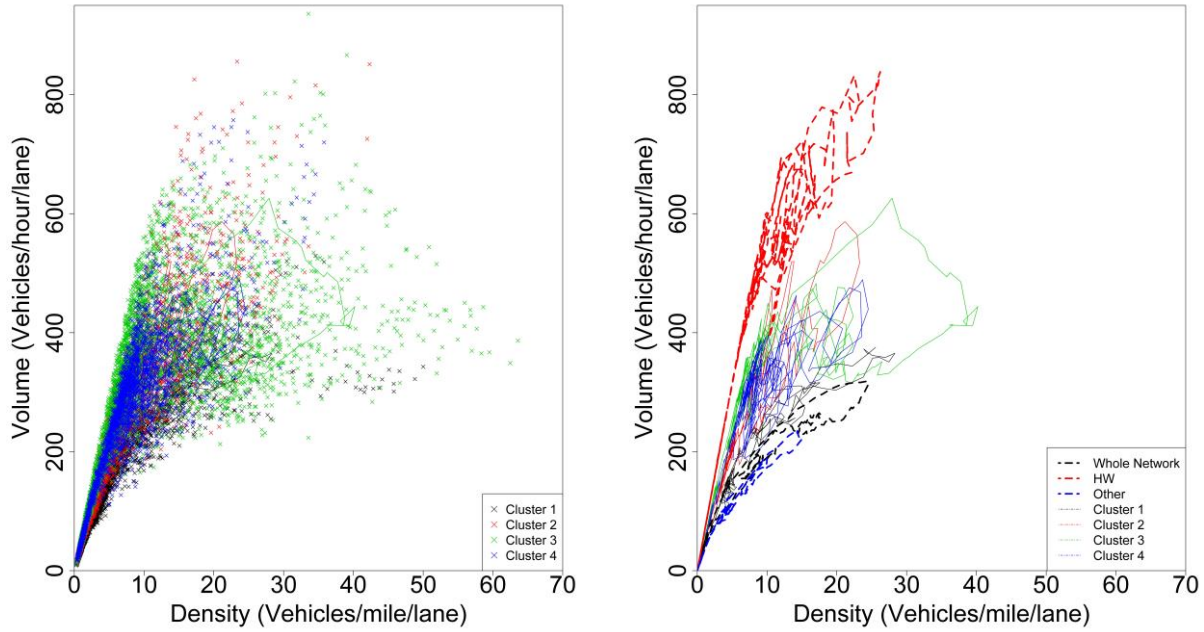
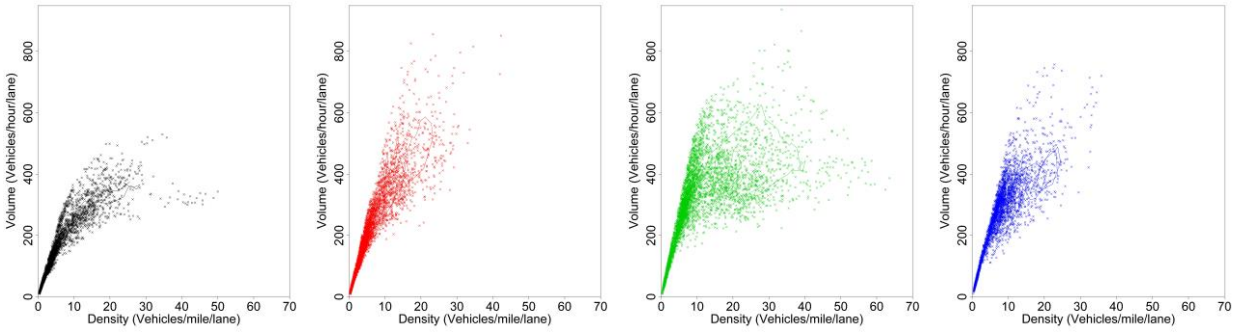


Figure 8-7 KM based a) NFDs for all the paths with four centers; b) centers of 4 clusters with NFDs of different road types



Cluster: 1 Cluster: 2 Cluster: 3 Cluster: 4

Figure 8-8 NFDs for all the paths by KM based clusters.

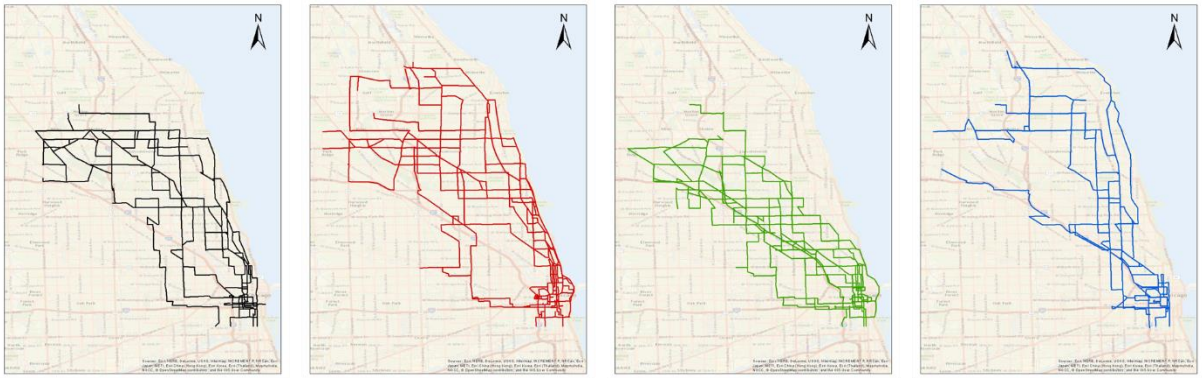


Figure 8-9 Spatial distribution of the paths by KM based clusters.

Results based on KM clustering are presented in Figure 8-7 through Figure 8-9. As in the case of DTW clustering, time series of velocities were used to identify the path clusters. Again, four significantly different clusters were obtained with very similar characteristics. However, the spatial pattern that emerged was more distinguishable when compared to the spatial patterns obtained from the DTW clustering.

8.3.3. TDTC clustering

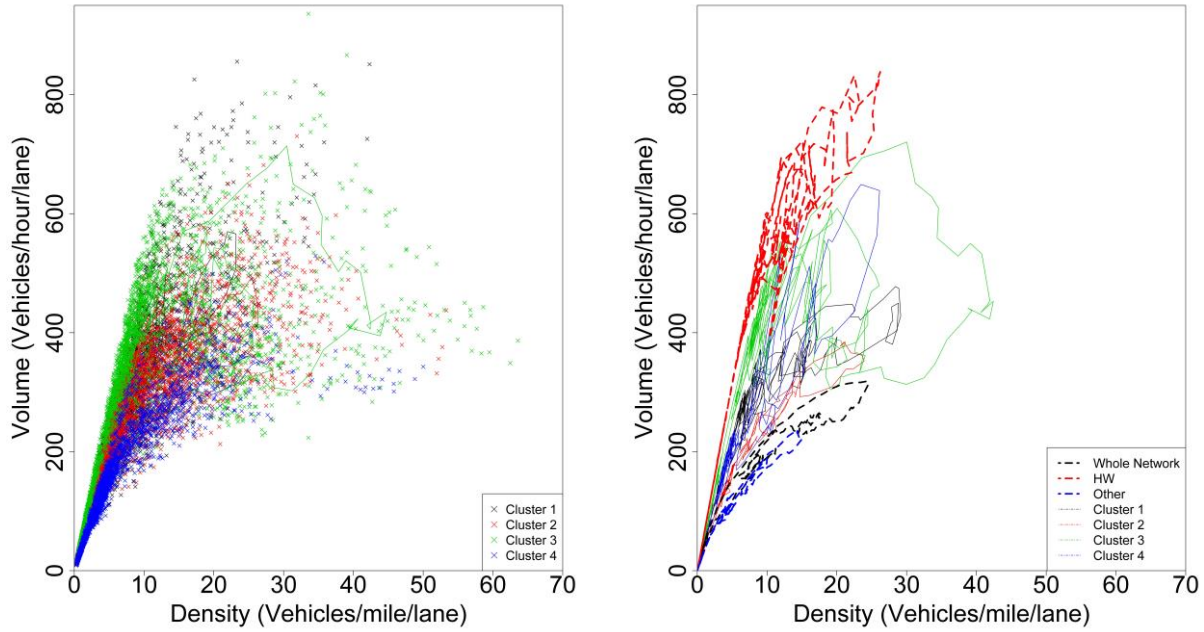
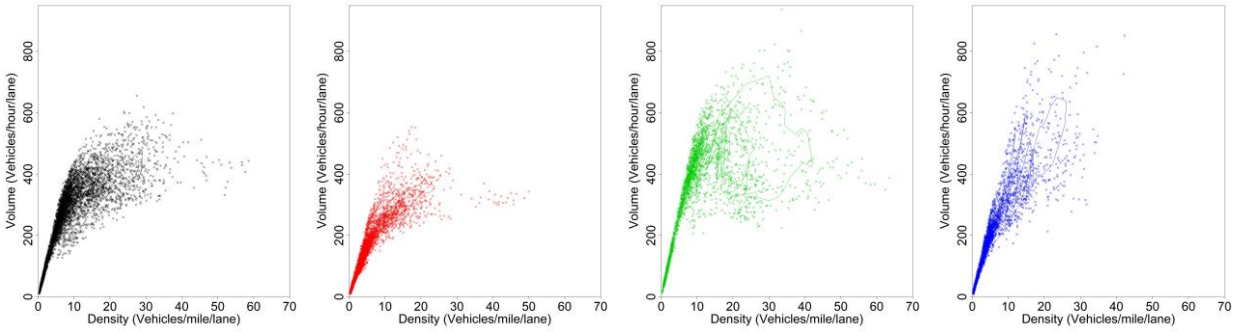


Figure 8-10 TDTC based a) NFDs for all the paths with four centers; b) centers of 4 clusters with NFDs of different road types



Cluster: 1 Cluster: 2 Cluster: 3 Cluster: 4

Figure 8-11 NFDs for all the paths by TDTC based clusters.

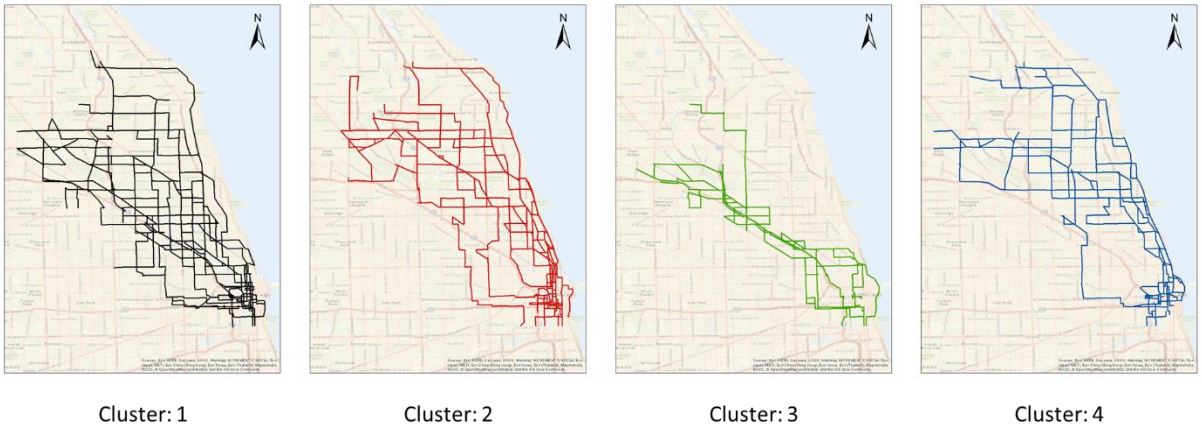


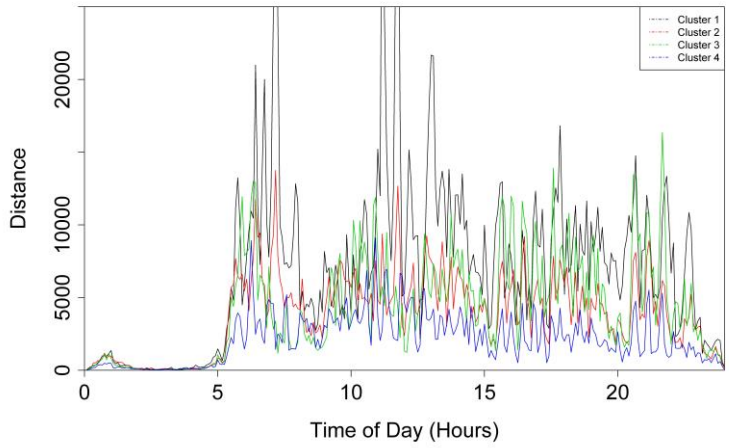
Figure 8-12 Spatial distribution of the paths by TDTC based clusters.

Results based on TDTC clustering are presented in Figure 8-10 through Figure 8-12. TDTC clustering produces results which are similar to the results obtained from DTW and KM clustering methods. However, as it can be observed in Figure 8-12 the spatial patterns of the clusters are even more prominent. Cluster 1 consists of, primarily, arterial roads with no or very few highway links. Cluster 2 and 4 are similar in spatial patterns but differ in their PFDs. Cluster 2 is characterized by a higher density and lower flow as compared to cluster 4. Cluster 3 comprises paths consisting of, predominantly, freeway links.

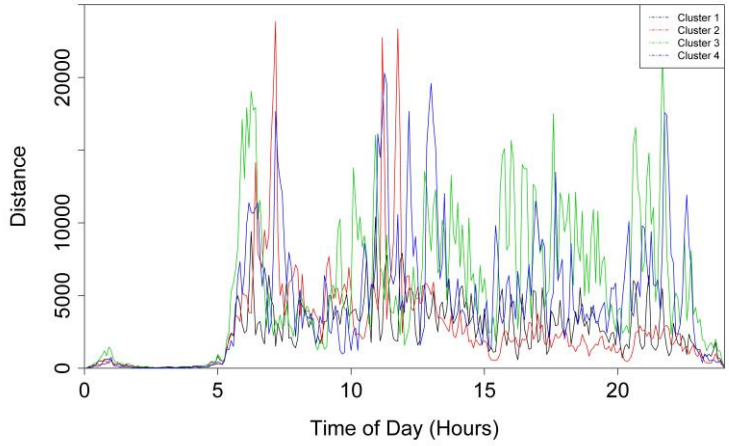
8.4. Comparison of clustering techniques

The three clustering techniques were compared based on variations in within-cluster distances.

a)


Cluster 1
Cluster 2
Cluster 3
Cluster 4
Distance
0
5000
10000
20000
0
5
10
15
20
Time of Day (Hours)

b)


Cluster 1
Cluster 2
Cluster 3
Cluster 4
Distance
0
5000
10000
20000
0
5
10
15
20
Time of Day (Hours)

c)

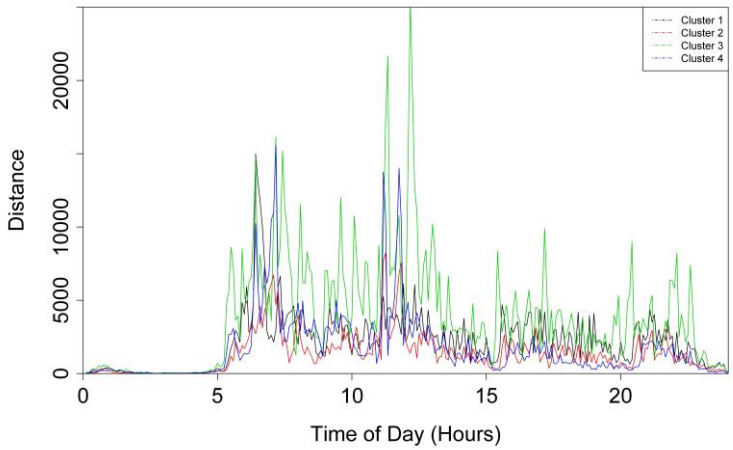


Figure 8-13 Variation among the clusters based on a) DTW; b) KM; c) TDTC method.

Presented in Figure 8-13 are the curves illustrating the within-cluster variations. For instance, Figure 8-13-a shows the variation within each of the 4 clusters based on the DTW method. The horizontal axis is the time of day, and the vertical axis is the mean distance from the cluster's mean value at the corresponding time instance. The average within-cluster variation and the overall average per clustering technique for the entire day are presented in Table 8-1. It can be observed that the within-cluster variations are the smallest for the TDTC method. Therefore, the TDTC can be regarded as a reliable technique to cluster NFDs or PFDs.

Table 8-1 Summary of variations within clusters and clustering technique.

	Cluster1	Cluster2	Cluster3	Cluster4	Average
DTW	6982	3871	4273	2160	4322
KM	2550	2905	5523	4625	3901
TDTC	2103	1361	3838	1728	2258

8.5. Common findings among clustering techniques

Emerging spatial patterns prompted further examination of the robustness of the clusters. Accordingly, paths which were clustered together with all the clustering techniques were identified.

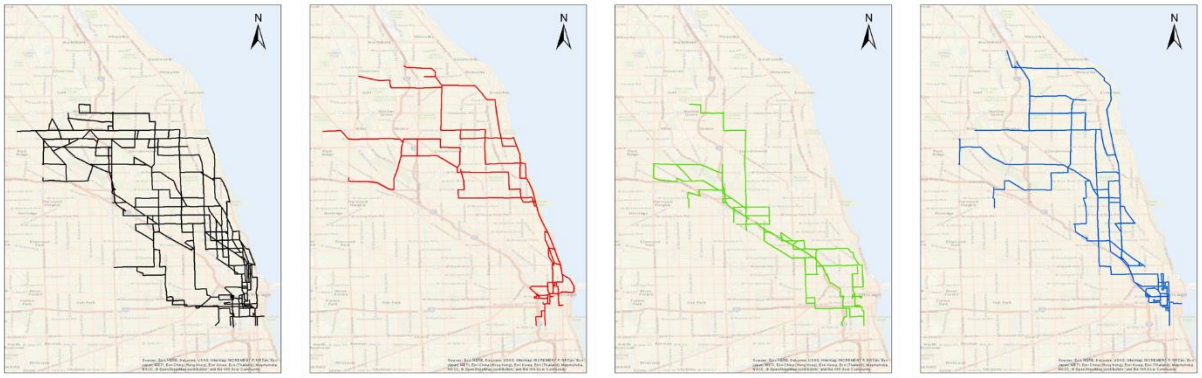


Figure 8-14 Spatial distribution of the overlapping path groups

Figure 8-14 represents 4 groups of paths that were found to emerge thorough any of the three techniques. However, the uniqueness of these groups is very evident. To summarize, cluster 1 represents arterial roads-based paths, cluster 2 - paths along the Lakeshore Drive highway, cluster3 - freeways, while cluster 4 paths comprise of mixed facility type - arterial and freeway - links.

8.6. Conceptual Applications: NFD (MFD) for control strategies design

With increasing research focus on NFDs, certain control and management policies upon NFDs have also been proposed. Geroliminis and Levinson showed that standard economic models of marginal cost could not accurately describe traffic congestion in time-dependent networks and suggested an equilibrium solution for a congested network in the no-toll case. The authors also devised a dynamic pricing model for morning commute as a combination of Vickrey's bottleneck model and NFDs (*86*). Similar work has been done by Arnott, R. and Fosgerau and Small. However, when applying the NFDs, the whole network was assumed to be homogenous (*89; 166*).

8.6.1. Local-wise Real-Time Traffic Control

A practical control technique could potentially involve traffic signals for the endogenous demand and pricing for the exogenous demand inflow. Practically, with historical and real-time sensor data, it would be feasible to adopt an anticipatory strategy for real-time prediction and then evaluate whether the proposed control policy would organize traffic into its optimal state in the near future, with respect to historically steady NFD.

Most real-time control strategies are focused on a single control method either pricing or traffic control or parking (*80; 167*). The effectiveness of a strategy will vary depending on the circumstances and type of facility, i.e., a small-scale network without endogenous demand versus an area with both endogenous and exogenous demand. The former could characterize arterial corridor operations, where almost the entire demand is externally generated. The latter usually

indicates a traffic high-generating area which also receives the demand from the neighboring areas such as CBD. Wu et al. conducted an empirical analysis based on arterial fundamental diagrams, focusing on the aggregated features of a network's traffic state (*90*). Conversely, small-scale network empirical studies instead of observing the macroscopic traffic state focused on the queue spillover effects with particular control policies (*81; 168*).

However, with respect to exogenous demand in an already congested area, the optimal strategy might be restricting the inflow from the outside. This might require changing the chosen mode of transportation, especially if traveling by car. On the other hand, for the internally generated demand, route or departure time guidance along with advanced signal control strategies might prove to be the optimal choice.

To accomplish this goal, an integrated control strategy with signal control for endogenous demand and pricing control for exogenous demand are designed for real-time networks.

In lieu of the work presented in this chapter, a potential control policy could consider the paths in a cluster as a single reservoir and accordingly control inflows (queuing behavior) as well as its corresponding outflows. The concept is aimed at decreasing the inflows in regions with high destination densities (unstable flow data points of an NFD) and managing the accumulation to maintain the flow at its maximum.

8.7. Corridor-wise Real-Time Traffic Control

The implementation of adaptive traffic signal control, dynamic lane assignment, and other corridor-based control strategies assume recognizing which operational conditions are most unfavorable.

Similarly, identification of high-demand corridors is dependent on the number of vehicles in the network (demand) and their respective chosen paths or trajectories. Therefore, the paths or trajectories (or sub-trajectories) with high demand can be regarded as the demand corridors. This analysis identified four clusters, consisting of paths that share similar NFD features, and thus are preferable to be regarded as a single entity when designing signal control policies for the paths within the same cluster.

8.8. Conclusions

In this chapter, the transportation system's user-defined paths were studied to identify the characteristics associated with those paths. In an attempt to study a network at a disaggregated level, researchers have been dividing large-scale networks into regional zones. Here, the network was studied at the path level. To identify popular paths, simulation-based vehicle trajectories were utilized. A popular path is defined as the path (or sub-path) shared by a considerable number of travelers given spatial and temporal constraints. To determine a set of popular paths a data-driven trajectories clustering method was adopted. Next, Path Fundamental Diagrams (PFDs) are obtained for each of the identified popular paths. PDFs are then clustered using three different methods; one of them is a novel method developed in this study. A comparison study shows that the new clustering technique, a two-dimensional time correlation (TDTC) method outperforms the other conventional methods. However, a certain level of consensus was found among all three clustering methods. Four different clusters were obtained from all three methods. An exciting outcome of the clustering was their spatial signature. Even though geolocations of the paths were not considered in the process, a spatial pattern emerged.

The study puts forward a new concept of disaggregating a transportation network based on user-defined paths. This method is different from conventional grid-based clustering. Based on this disaggregating concept, a new clustering algorithm was developed and demonstrated to outperform the other explored clustering methods.

Most noteworthy features presented in this chapter are as follows:

- Paths based on drivers' or users' route choice behavior are identified and clustered
- Relationships between traffic flow parameters are determined on a path level
- Fundamental diagrams hold at (user-defined) path level
- A new theoretical basis for network partitioning was suggested – a path or route choice behavior-based clustering, which incorporates a time of day correlations.

Reported results demonstrate that even networks with similar NFDs may require different management strategies depending on the time of day traffic patterns. Hence, spatial and temporal profiling of a large-scale system is needed, that could help planners and practitioners design dynamic control and flow management strategies. In summary, the study explains how to divide a regional network into sub-networks, based on demand patterns dynamics. Incorporating the time of day correlations allows for control/management strategies to be dynamic. An online application of such study could, potentially, be deployed for real-time traffic flow control.

CHAPTER 9. TRAVEL TIME DISTRIBUTION ALONG USER-DEFINED PATHS

Travel time between two points is a random variable with non-stationary distribution. It is stochastic in that a single value cannot be associated with it; only a range can be given with confidence. It is non-stationary because the associated probability distribution changes with time as well. Consequently, to determine travel time on a facility, observations covering an entire day over multiple days and seasons are required. A simple distribution obtained from data for a particular time of day say noon peak period, is not enough.

In the previous chapter, the transportation system was characterized based on user-defined paths. To understand the system at the same level, in this chapter, the assessment of travel time reliability is the main premise.

Availability of such a significant amount of observations is a challenge in itself. Several data sources have been explored and utilized to conduct such studies. The data sources include, for example, GPS, probe vehicles, computer simulation models, and wireless signatures. Observations from the sources, located on the same facility, can be combined to develop a richer data set in a consensus format. With the rapid development of information and communication technologies (ICT), equipping automobiles with wireless communication capabilities are expected to be the next frontier for automotive evolution. Connected vehicles are proactive, cooperative, well-informed, and coordinated, and will pave the way for supporting various applications for road safety (e.g., collision detection, lane change warning, and cooperative merging), smart and green transportation (e.g., traffic signal control, intelligent traffic scheduling, and fleet management), and location-

based services (e.g., point of interest and route optimization). Connected vehicles offer valuable information regarding the performance and state of the vehicle and its trajectory, including time-stamped position, direction, speed, routing, and driving style and preferences. Connected vehicles can transmit their location and progression-related information in a high-resolution format and at a faster rate than any other currently available technology. Vehicle to vehicle (V2V) and vehicle to infrastructure (V2I) communication along with connected and automated vehicles (CAVs) makes this information readily available. Trajectories provide in-depth information about users' travel experience in a transportation system. Moreover, trajectory-based measurements are more accurate than those based on link (aggregated) information.

On the other hand, behavior studies have consistently shown the importance of travel time reliability to user satisfaction with the experience delivered by a transportation network, as well as to the travel and activity choices that individuals make (*169*). Travel time reliability also has a direct economic impact on the decisions of logistics firms and delivery vehicles, and the service levels ultimately experienced by shippers, firms, and consumers. For example, Fosgerau et al. (*169*) assume a deterministic model of travel time distribution for a given departure time of the day to model the utility of reliability in scheduling. Authors illustrate that the standardized travel time distribution is independent of time of day, but not the mean and standard deviation. With this in consideration the developed model has mean and variance as a function of departure time and the standardized trip duration distribution independent of the departure time.

A critical question that arises from the perspective of a given user (i.e., a traveler or good shipping company) pertains to the variability of travel times along specific paths contemplated by the user (*170*). Related to that is the problem of finding paths that are in some way optimal (*171*) or meet

specific reliability targets (*172*) when the network travel times are stochastic with non-stationary distributions (*173*). Accordingly, there is a need to study the travel time as a function of time and day. The mere addition of expected travel time over the links to obtain a path's travel time is not sufficient.

Underlying all measures of travel time reliability is the variability of travel times, experienced along individual links, by various modes, at junctions and intermodal terminals, through the individual trajectories of travelers and goods. The importance of trajectories as the complete representation of the travel experience has been recognized recently in the literature (*102*), particularly in project L04 under the second strategic highway research program (SHRP2) (*174*).

Most early work on pathfinding with stochastic travel times assumed that travel times are independent random variables; a variation on one link at a given time is treated as an independent process from a variation on adjacent links, or at different times. The main factors known to affect travel time variation, such as congestion and weather, tend to act on multiple links simultaneously, and their impact tends to linger long after the event itself may have cleared. This results in varying degrees of correlation in the travel times observed on different links in different time periods. While correlation tends to be stronger along contiguous or adjacent links, it is not limited to these situations, as recently shown by Zockaie et al. (*109*). Accordingly, the study seeks to develop a methodology to incorporate link travel time correlation. Estimating general correlation patterns is challenging because it requires a much more significant number of observations than estimating the first or second moments of these distributions. Also, obtaining the path-level distributions requires convolving the link-level distributions, a process that typically do not have closed-form analytical solutions, and for which numerical integration techniques may not always converge. For

this reason, statistical sampling methods, such as Monte Carlo techniques, have been used effectively for that purpose in a variety of areas. For example, calculation of sample likelihoods for maximum likelihood estimation of discrete choice models with general correlation patterns, such as multinomial probit or mixed logit, is commonly performed with statistical sampling techniques using both quasi-random methods such as Monte Carlo or pseudo-random techniques such as Halton sequences.

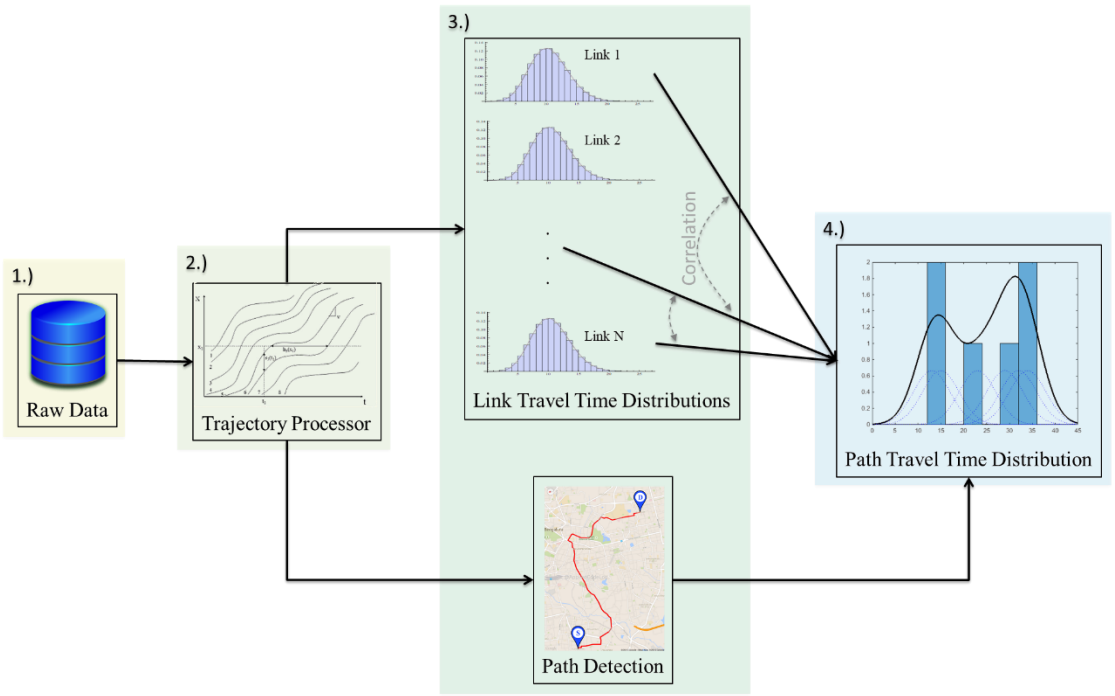


Figure 9-1 Methodology to estimate path travel time distribution.

In the interest of addressing the aforementioned subjects, the study aims at 1) using vehicle trajectory data to identify the paths traversed by drivers, recognizing that these paths are continuous but may comprise of different facility types, and 2) using the same data set to develop

the methods to successfully synthesize and replicate traveler-based distributions of travel time along the identified paths. The methods shall formulate and develop the path-level distributions as convolutions of link-level travel times that follow general distributions (i.e., recognizing spatial-temporal correlations). Sequence steps need for such a study are presented in Figure 9-1 above.

9.1. Methodology framework

As is shown in Figure 9-1 above, the sequential framework consists of four modules: a) raw data, b) trajectory processor, c) link-level information library, and d) path travel time distribution.

a) **Raw data**: Real data retrieved from infield observations are frequently contaminated with noise, missing values, erroneous measurements. In data-driven studies preliminary step is to clean and scrub the data to make it readily available for investigation. After cleaning data is reformatted as per analyst's desire.

b) **Trajectory processor**: Cleaned and reformatted trajectories obtained from the previous step are processed to create a library containing travel time distribution for links individually. At step, also the path, i.e., a set of consecutive links is identified between a given origin-destination (O-D) pair.

c) **Link level information library**: For the identified paths and with the library, links comprising the paths are a subset of the library, and correlation relationship among the paths is identified.

d) **Path travel time distribution**: Finally, with all the information retrieved from the previous steps, convolving integral with correlated random variables is solved to determine the underlying travel time distribution.

9.2. Existing data sources and types

The study aims to develop solution methods with the real trajectories data which calls for a need of adequate data source. Popular data sources found for such a study include 1) fixed sensor, 2) floating car data, and 3) GPS information from equipped vehicles. Fixed sensors are limited in the coverage and the information they can provide. In general, they provide occupancy and speed. This leaves the analyst to estimate the flow and density. Floating car data (FCD), provide information such as speed, location, the direction of travel from active mobile phones in the vehicles. This provides better coverage and more information as compared to fixed sensors. In addition, to improved data quality, FCD also eliminate a need for additional hardware installation or any approximations. To further enhance the quality of the information, data from the vehicles equipped with GPS are being used.

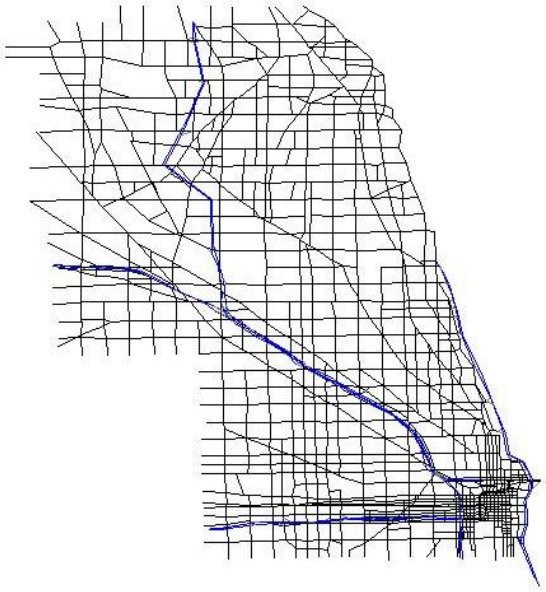
Sources used in the study:

- Simulated trajectory data
- Google Maps "*Distance matrix*" API
- TomTom trajectory data

9.2.1. Simulated Trajectory Data

Simulated vehicle trajectories were used to conduct the analysis. Trajectories were obtained from the Chicago City network simulation model, where operational conditions were established as the clear weather with no precipitation and normal visibility.

Chicago testbed was extracted from the entire Chicago Metropolitan Area network to enhance the estimation and prediction performance during the implementation procedure. This testbed represents a section of the network previously described in Chapter 4 and shown in Figure 9-2. Detailed calibration and other characteristics are out of the scope of this study and can be found in the previous work (*45; 46*). Data contains information for vehicle regarding its entry, travel time and exit time for each link it traveled on.



Number of nodes	1,578
Number of links	4,805
Number of vehicles	805,275
Demand duration (h)	24

Figure 9-2 Schematic Diagram of Chicago network.

9.2.2. Google Maps "Distance Matrix" API

Google provides an application programming interface (API) called "Distant Matrix". Utilizing the API, it is possible to retrieve a temporal profile of travel time between two points. For the purpose of the study, the API was used to retrieve travel time on a 4-mile stretch of West Peterson Avenue, Chicago which connects I-94 Freeway on west and Lakeshore Drive Highway on the east side (Figure 9-3). It has 8 signalized intersections with spacing ranging between 0.17 and 1 mile and average spacing being 0.56 miles. Travel time was collected over several days, over 20 days so far, in one-minute intervals. Different from GPS data, this data provides only one data point for a timestamp and an OD. Please note that the data is being collected continuously to make an extensive dataset.

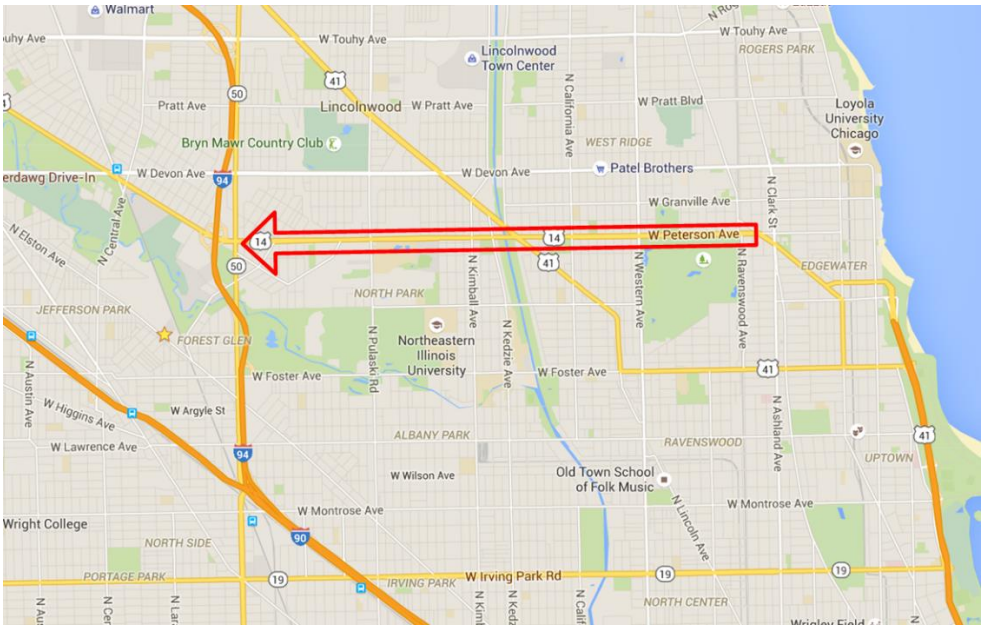


Figure 9-3 West Peterson Avenue, Chicago (Source: Google Maps)

Figure 9-4 and Figure 9-5 the temporal profile of the collected travel time by day of the week. Profiles clearly capture the morning, mid-day and evening peaks over the weekdays. Over the weekend, due to lower demand, the travel time is observed to reduce as compared to that on weekdays.

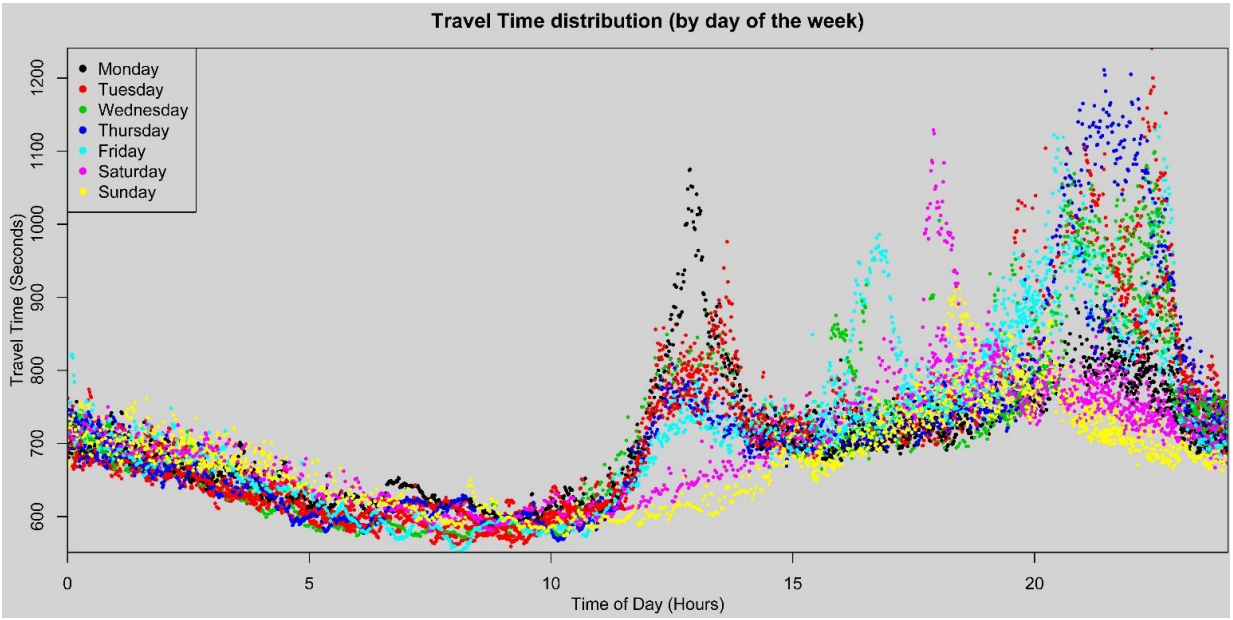
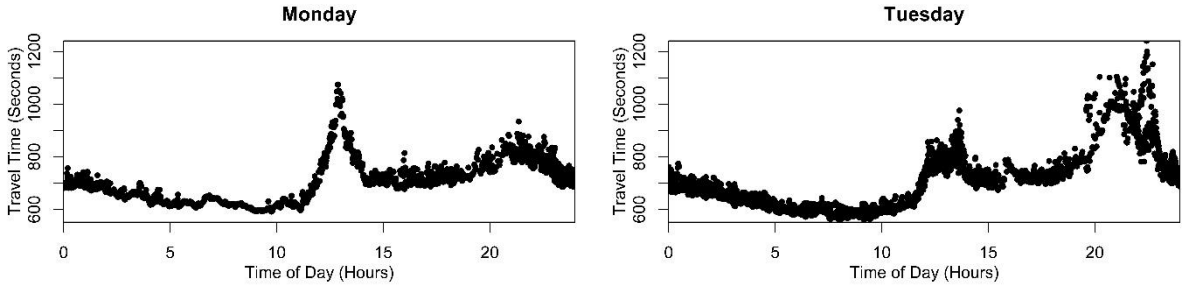


Figure 9-4 Temporal profile of the collected data by day of the week



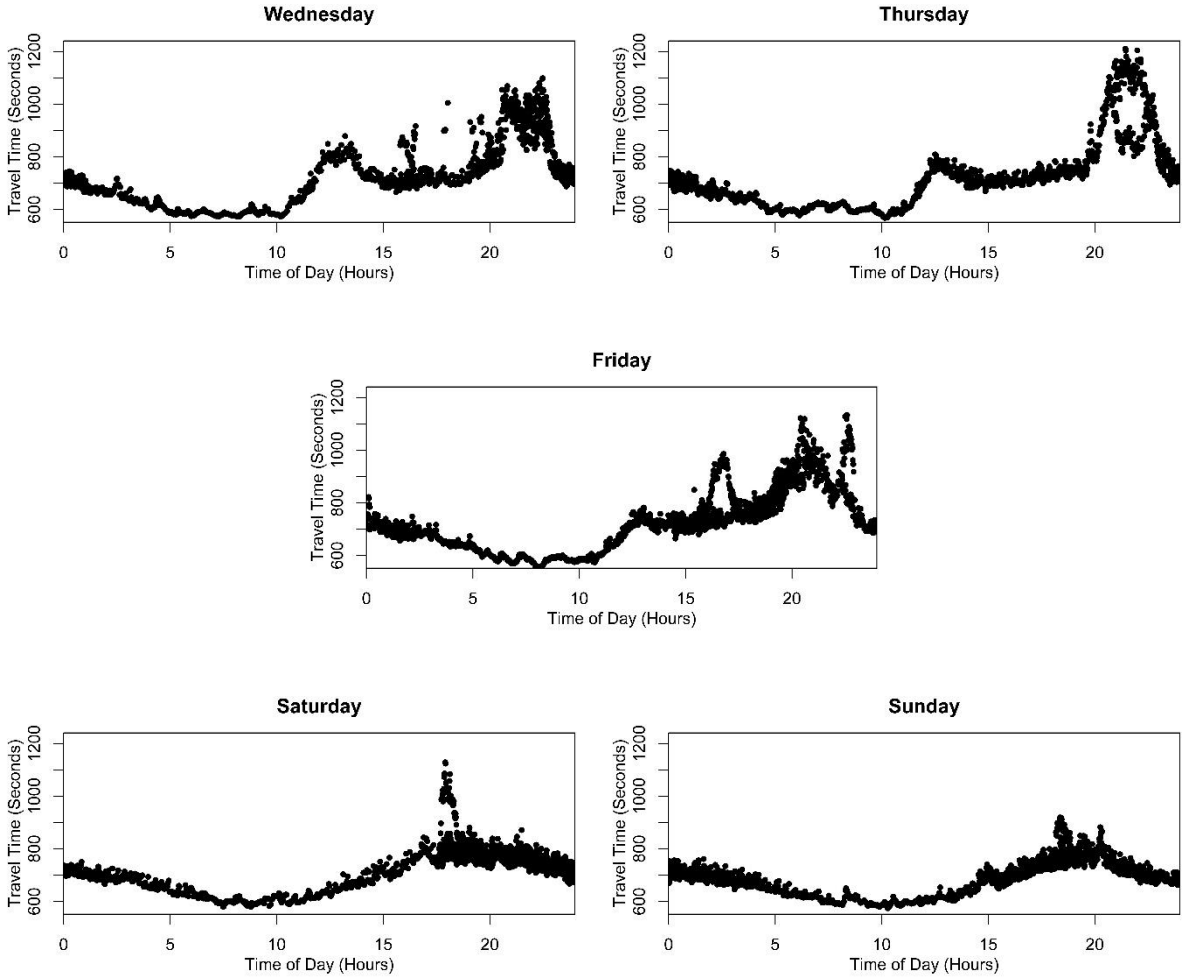


Figure 9-5 Temporal profile of the collected data by day of the week

From the figures, it can be observed that before 10 AM all days have a similar trend in the travel time. The difference in the trend starts to emerge after 10 AM. Between 10 AM and 3 PM, there is a peak for weekdays, where the peak is the highest for Monday. Over the weekend no peak is observed in the time period. The second peak in the travel time distribution is observed in the evenings of the weekdays. This peak was the lowest for Mondays, contrasting to morning peak

trend. A third peak was observed on the Fridays around 4-5 PM. This can be due to people leaving work early on the day. For the weekends, the trend showed no significant peak.

9.2.3. TomTom Trajectory

The acquired trajectory data pertains to the data provided by TomTom for an SHRP2 study. Following is a brief description borrowed from the SHRP2 study report. Detailed information can be found in the original report. (*7*)

The observed probe data set was collected through TomTom portable navigation devices (PNDs); whenever a vehicle containing this device moves along a road network, a trace/log is produced that contains information related to location and time for that vehicle. Data were obtained from TomTom, which covered approximately 1,100 square miles within the New York metropolitan area. The coverage area is shown in Figure 9-6. The raw data contained information for 16 full days, spanning from Sunday, May 2, to Sunday, May 16, 2010. There were approximately 2.5 million probe data records provided for each day. For privacy protection reasons, the information provider randomized some of the data records. Records related to the first and last 10 segments of a unique probe ID were aggregated into one record, and time stamps were randomized by ±6 min. Figure 9-7 shows a snapshot of the data contained in the raw probe data.

Figure 9-6 TomTom data coverage area

The first column is the probe identification number (ID) and is unique for each navigation device. The second column is the section the vehicle is traveling along. Each section is geospatially referenced and is contained in a geographic information system (GIS) map. The unique ID numbers for each section contained in the probe data correspond to the section IDs in the GIS database. The third column contains the ingress time (in milliseconds) for each section. This is a time stamp for when the vehicle enters each section. The time stamp is referenced to provide an instant in the time measuring from the UTC epoch of January 1, 1970. The fourth column is the travel time in seconds that the vehicle takes to traverse each section of the network. The fifth column provides the section length traversed in feet.

```
64555,startCut,1272751279861,60.62,1992.8
64555,UN2_4361986,1272751340861,3.05,133.5
64555,UN2_4409098,1272751343861,1.81,94.2
64555,UN2_4404568,1272751345861,0.72,38.7
64555,UN2_4361997,1272751346861,4.26,236.5
64555,UN2_4362106,1272751350861,1.60,93.2
64555,UN2_4362105,1272751351861,2.69,157.2
64555,UN2_4358640,1272751354861,1.74,101.7
64555,UN2_4424651,1272751356861,6.37,371.7
64555,UN2_4424650,1272751362861,9.58,209.6
64555,UN2_4424080,1272751372861,4.82,145.0
64555,UN2_4424081,1272751377861,2.59,96.8
64555,UN2_4362014,1272751379861,6.69,250.0
64555,UN2_4409103,1272751386861,6.23,232.6
```

Figure 9-7 Raw probe data

In summary, the raw data contain individual records by section traversed by each probe within a 24-h period. It, therefore, means that multiple trips taken by an owner of the PND are contained within the probe records for one day and to distinguish between the trips the data had to be processed to produce the trajectory traces related to separate trips for the same probe ID.

1.1.1.1 Processing of the raw data:

The raw data was processed to determine the a) link and b) route level travel time distributions.

a.) **Link level travel time distribution**:

Link level information was retrieved from the trajectories. All the trajectories which overlapped significantly with the link into consideration were accounted towards the travel time distribution of the link. To begin the analysis at this level, the whole city of New York was divided into pre-defined links. And then the trajectories were mapped to the links. Not all the trajectories swept the entire link into consideration. For the case when only a part of the trajectory was overlapping the link, a weight (<1) was assigned. The weight of trajectory, ϕ_i, towards the travel time distribution of link l is calculated as:

$$w_{\phi_i}^{l} = \frac{overlap\ of\ \phi_i\ with\ l}{length\ of\ l} \quad (9\text{-}1)$$

Hence, weight travel time rate distribution was determined for each link. The trajectories with weight greater than 0.5 were taken into consideration. The developed travel time distribution forms the library of the of link level time distribution as presented in Figure 9-1. The library is referred to retrieve link level travel times to solve the convoluting integral of correlated random variables. The problem statement and process are described in section 9.3below.

b.) **Route level travel time distribution**:
A route is defined as a set of consecutive links. Information at this level of analysis forms the ground truth. Travel time distribution of a route into consideration was determined through the trajectories overlapping with the entire route.

9.3. Numerical Solution Method

The fundamental problem is to formulate and solve a convoluting integral where random variables are correlated. Moreover, the random variables may have distribution types that are different from each other. In this section, an integral that includes these aspects is formulated first. Model for correlation is yet an ongoing investigation.

A path, P, consists of N consecutive links $\in \{l_1, l_2, l_3, \dots, l_N\}$. Travel time on link l_i, $\forall i \in \{1,2,3, \dots, N\}$, is denoted by θ_i. Then the distribution of travel time on the path P, $\Theta_P(\mathrm{T})$, is

obtained from solving the integrals presented below. First link travel time are assumed to be independent and then dependent. In the dependent case, the model is made sophisticated by introducing non0stationary and operational conditions.

9.3.1. Convoluting Integral with Independent Random Variables

Convoluting integral with an independent random variable is solved using the method of induction. That is, the integral is solved with two random variables at the time where the result from the previous step forms a new random variable for the next step.

$$\Theta_P(\mathrm{T}) =$$

$$\int_0^\infty \dots \left(\int_0^\infty \left(\int_0^\infty \theta_1(\tau_1) \,.\, \theta_2(\tau_2 - \tau_1) d(\tau_1) \right) . \theta_2(\tau_3 - \tau_1 - \tau_2)\, d(\tau_2) \right) \dots \theta_N \left(T - \sum_{j=1}^{N-1} \tau_j \right) d(\tau_{N-1}) \qquad (9\text{-}2)$$

9.3.2. Convoluting Integral with Dependent Random Variables

When the correlation is assumed then the distribution of travel time on the path P, $\Theta_P^o(\mathrm{T})$, of order o is obtained from solving the following integral:

$$\Theta_P^o(\mathrm{T}) =$$

$$\int_0^\infty \int_0^\infty \dots \int_0^\infty \theta_1(\tau_1).\theta_2(\tau_2-\tau_1).\theta_2(\tau_3-\tau_1-\tau_2)\ \dots\ \theta_N\left(T - \sum_{j=1}^{N-1}\tau_j\right) d(\tau_{N-1}) \dots d(\tau_2) d(\tau_1) \qquad (9\text{-}3)$$

In the formulation, the order $o \in \mathbb{Z}^{\geq}$ is the maximum number of upstream links that are considered to be correlated. Accordingly, θ_i is assumed to be correlated with the previous $n = \min\{i, o\}$ links. In equation (9-3), probability distribution type of θ_i's is not restricted to any specific from. Furthermore, links within a path can have multiple distribution types of travel time.

There are vehicles that do not travel the entire stretch of the path, but only a portion. Before they turn away from the path, these vehicles may slow down and thus impact the continuing vehicles. Hence, accounting for these vehicles in the same way as the continuing vehicles will give an incorrect estimate of path travel time. To address this issue, continuing vehicles to the next link on the path from the current link are given more weight to compute link travel time distribution while solving the integral in equation (9-3).

The outcome of the integral is a distribution of travel time which will still be needed to fit into an analytical form. The first attempt is to fit a standard distribution type, like log-normal distribution. Another approach is to estimate a Fourier series. A decision criterion other that RMSE is needed because it was found that such estimation techniques are inaccurate around the mode values of the distributions.

Integral presented in equation (9-3) is stationary, to add a temporal dimension to it the integral is modified to take the following form:

$$\Theta_P^{o,t'}(\mathrm{T}) =$$

$$\int_0^{\infty}\int_0^{\infty}\dots\int_0^{\infty} \theta_1^{t'}(\tau_1).\theta_2^{t'}(\tau_2-\tau_1).\theta_3^{t'}(\tau_3-\tau_1-\tau_2)\dots\theta_{N-1}^{t'}\left(\tau_{N-1}-\sum_{j=1}^{N-2}\tau_j\right).\theta_N^{t'}\left(T - \sum_{j=1}^{N-1}\tau_j\right) d(\tau_{N-1})\dots d(\tau_2)\,d(\tau_1) \quad (9\text{-}4)$$

In equation (9-4), t' is prevailing time interval which simply can be four time-periods: 1) Early morning, 2) Mid-day, 3) evening, and 4) late evening to midnight. t' can also be finer time interval bins of one hour. Former time intervals save on computational efforts and the later gives a precise estimation. However, for the later time intervals, there might not be enough data for estimation. Furthermore, there is an effect from operational conditions (OCs) as well. Accordingly, weather or seasonal attribute to represent different OCs is also added to equation (9-4).

$$\Theta_P^{o,\zeta,t'}(\mathrm{T}) =$$

$$\int_0^{\infty}\int_0^{\infty}\dots\int_0^{\infty}\theta_1^{\zeta,t'}(\tau_1).\theta_2^{\zeta,t'}(\tau_2-\tau_1).\theta_3^{\zeta,t'}(\tau_3-\tau_1-\tau_2)\ \dots\ \theta_{N-1}^{\zeta,t'}\left(\tau_{N-1}-\sum_{j=1}^{N-2}\tau_j\right).\theta_N^{\zeta,t'}\left(T-\sum_{j=1}^{N-1}\tau_j\right)d(\tau_{N-1})\dots d(\tau_2)\,d(\tau_1) \quad (9\text{-}5)$$

In equation (9-5), ζ is operational condition indicator. It can be divided in as many categories as a need, but it will increase computational efforts. If too few categories are used, it can cost the efficiency of the method. Hence, there is a trade-off here as well. An optimal balance will be found through conducting a sensitivity analysis.

9.3.3. Solution Method

In general, the convolution may not have a closed form particularly if the distributions are different. In such cases, numerical approximate solution techniques are established. For instance, let there be a convoluting integral with two correlated random variables. In this case, one simple technique would be to condition on them sequentially. This is to say that condition on one of the random variables and then un-condition on it. For example, if we want to find the distribution of $Z = X + Y$, then the first condition on y, i.e., $P(Z \leq z) = P(X \leq z - y|Y = y)$ and then un-condition on y. Another more complex method is to find a characteristic function (or the moment generating function as found in the studies in (*15*) and (*16*)) of Z as the product of characteristic

functions of X and Y and then invert that characteristic function. However, such methods may not scale well at a large network level. To achieve computational efficiency in the estimation method, commonly, a sampling technique as in Monte Carlo simulation-based techniques are exercised. In an overview of such techniques, given a distributions of each link travel times which form a path and the correlation among them, conditional probabilities are to be established. Based on thus formed probabilities, samples are estimated to form a vector of realizations of the path travel times. If the correlation is ignored, i.e., samples are drawn independently, the result would be the same as to the solution of a regular convolution integral that assumed no correlation among its random variables.

This study adopted a Markov Chain Monte Carlo simulation-based technique to produce a non-stationary, non-parametric solution for the link travel time distribution based on the individual link travel time distributions. Two foundational steps in the technique are:

- Estimate link travel time and correlation among them.
- Solve convoluting integral with correlated random variables to convolve the link travel times together to compute path travel time.

In what follows, four sampling techniques were used to reconstruct path travel times from corresponding link travel times:

1. Link travel times were assumed to be independent
2. Quasi-time dependency was introduced
3. True time dependency was introduced
4. True time dependency with correlated travel time was introduced

9.4. Independent Link Travel Times

At this step, link travel times were assumed to be independent, i.e., uncorrelated. This step is based on the following iterative procedure:

- One realized travel time was picked randomly for each of the links. A set of link travel times is constructed.
- Picked link travel times are summed together to estimate a path travel time.
- This process is reiterated multiple time to obtain a set of path travel time estimates.

9.5. Quasi-Time Dependent Link Travel Times

Here, the link travel times were assumed to be quasi-time dependent. The steps involved in this process are as follow:

- Whole time duration is binned in uniform 15-minutes intervals.
- A time bin is chosen randomly,
- A realized link travel time for each of the constituting links is picked from the chosen time bin to form a vector.
- Elements of the vector are summed to get a path travel time estimate for the time bin.
- The process is repeated multiple times.

9.6. Time-Dependent Link Travel Times

This step is based on the previous step of quasi-time dependent link travel times. In the previous step, link travel times from the same (departure time) bin, say the 10:30 AM to 10:45 AM bin,

were selected to produce a travel time for the path for that same bin. However, this is not true time-dependence, because the bin for a downstream link, in reality, may be different; for example, if the travel time from link 1 to link 5 is more than 15 minutes (technically 7.5minutes if selected a point at random), then one would be in a different time bin for the downstream link. Time-dependence is about recognizing that the arrival time at a downstream node depends on the travel time along link or sub-path to that node. This is likely to be even more of an issue for the shorter time bins (e.g., 2-minute). To incorporate the true time dependency, the time bins were adjusted based on the arrival time at the start node of the link.

9.7. Time-Dependent Correlated Link Travel Times

Building on the previous step, at this stage correlation among link travel time in incorporated. That is, traffic flow conditions as experienced by a user on the upstream link is accounted for. The rationale behind this incorporation is that just basing on the arrival time at a link does not sufficiently determines a user's experience on the route. To incorporate correlation, links traffic flow conditions were clustered to identify congested and uncongested regimes. Conditional on the upstream link's congestion level, downstream link's realized travel times were utilized to determine path's travel time.

9.8. Correlation Structure

As pointed out by the authors in (*106*), most of the literature assumes independence among link travel time. This assumption leads to biased results that are far off actual measurements. If the

correlation is ignored, the estimated trip/path travel time is not accurate and is typically underestimated (*105*). It is essential to capture and understand the correlation among the link travel times for studies that measure reliability in the system. Methodologies that incorporate correlation are readily applicable to the case where link travel times are independent, but the converse is not true. To model independence in such methodologies, merely setting the correlation matrix to zero would suffice as independence is a particular case of zero correlation. The challenge of summing correlated random variables often arises in cellular communication systems where a user preferred bandwidth is favored for efficient reception. Reference (*175*) is a study on calculating an approximate sum of lognormal RVs and has references on the same subject. However, in communication systems, link travel time is very short, and the networks are typically much denser than transportation networks, so a large number of paths may be available for close-to-optimal routing.

Recent work by Zockaie et al. (*109*) on link travel time correlation address this in a recent study. The study provides an algorithm to compute the correlation structure with a different order of neighboring links. The developed algorithm is applied to a stochastic network. Their methodology incorporates temporal (within a day) and spatial variability. The correlation model is as follows:

$$Cor(x_t, y_{t'}) = \frac{\sum_{i=1}^{I}\left[\left(x_t^i - \overline{x_t}\right)\left(y_{t'}^i - \overline{y_{t'}}\right)\right]}{\sqrt{\sum_{i=1}^{I}\left[\left(x_t^i - \overline{x_t}\right)^2\right]\sum_{i=1}^{I}\left[\left(y_{t'}^i - \overline{y_{t'}}\right)^2\right]}} \quad (9\text{-}6)$$

where x_t and $y_{t'}$ are travel times of vehicle i at time interval t and t' respectively and $t' = \{t-1,\ t,\ t+1\}$.

Further, on correlation structure, Rakha et al. in (*176*) clearly state that assuming independence of link travel-times gives wrong estimates of variability of travel time on a path. Estimating mean path travel time is a simple addition of mean link travel times but as claimed in reference (*110*) mean is not sufficient to estimate a system's reliability. Rakha et al. present multiple estimators of link travel time correlation. With given n, number of vehicles $\in V(s)$, on path s, path s composed of $m \in L(s)$ links, and t_{ij} being the travel time realized by a vehicle $i \in V(s)$ on link $j \in L(s)$. trip travel time for a vehicle i the is calculated as:

$$t_{it} = \sum_{j \in L(s)} t_{ij}\,, \forall i \in V(s) \qquad (9\text{-}7)$$

moreover, expected travel time along a link is computed as:

$$\bar{t}_j = E_i\{t_{ij}\} = \left(\sum_{i=1}^{n} t_{ij}\right)/n\,, \forall j \in L(s) \qquad (9\text{-}8)$$

Similarly, for a trip, the expected trip time is

$$\bar{t}_t = E_i\{t_{it}\} = E_i\left\{\sum_{j \in L(s)} t_{ij}\right\} = \sum_{j \in L(s)} \bar{t}_j \qquad (9\text{-}9)$$

Authors basing their research on other literature and own techniques give five methods to compute the variation of path travel time from composing links' travel times retrieved from the vehicles who traversed on them. Table 1 gives five estimators of variations developed by the authors.

Table 9-1 Travel time variability methods

Method	Symbol	Equation	
1	$\hat{\sigma}_{1t}^2$	$\left\{\sum_{i\in V(s)}\left(\sum_{j\in L(s)} t_{ij}^2\right)\right\}\Big/ n - \sum_{j\in L(s)} \bar{t}_j^2$	(9-10)
2	$\hat{\sigma}_{2t}^2$	$\frac{\bar{t}_t^2}{\sum_{j\in L(s)} \bar{t}_j^2}.\left\{\left(\sum_{i\in V(s)}\left(\sum_{j\in L(s)} t_{ij}^2\right)\right)\Big/ n - \sum_{j\in L(s)} \bar{t}_j^2\right\}$	(9-11)
3	$\hat{\sigma}_{3t}^2$	$\frac{\bar{t}_t^2}{m^2}.\left(\sum_{j\in L(s)} \frac{\sigma_j}{\bar{t}_j}\right)^2$	(9-12)
4	$\hat{\sigma}_{4t}^2$	$\left\{\bar{t}_t.Med_j(CV_j)\right\}^2$	(9-13)
5	$\hat{\sigma}_{5t}^2$	$\frac{\bar{t}_t^2}{4}(CV_{max} - CV_{min})$	(9-14)

$\hat{\sigma}_{1t}^2$ assumes independency of link travel times. $\hat{\sigma}_{2t}^2$ was developed based on the bounds on covariance as described in (*177*). $\hat{\sigma}_{3t}^2$ assumed the trip CoV is expected value of the realized CoV of the segments. CoV is coefficient of variation calculated as the standard deviation divided by the mean.

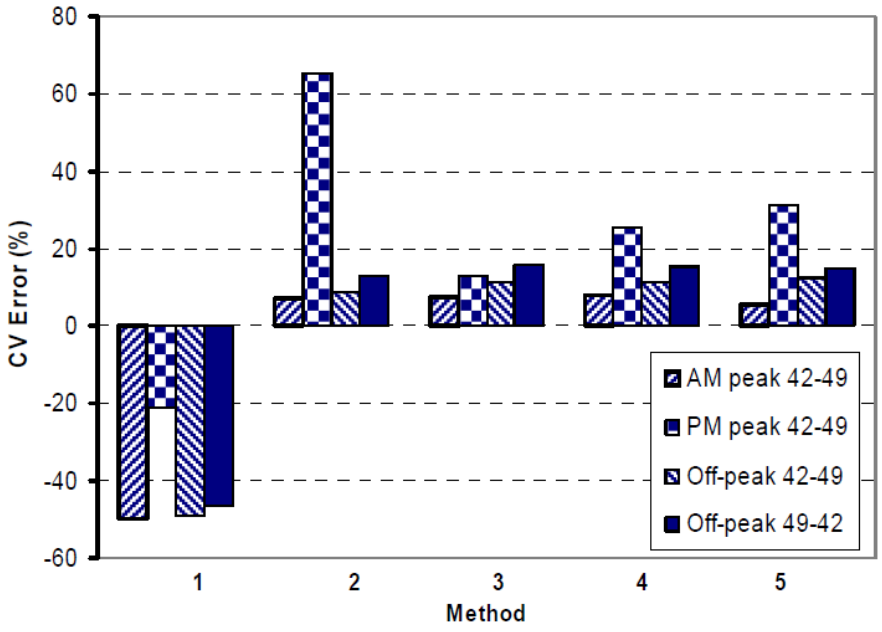


Figure 9-8 Comparison of the methods [Source (*176*)]

Figure 9-8 presents the comparison of the five methods. Authors found $\hat{\sigma}_{3t}^2$ to give the least error between the estimates and empirical data. However, they also insist on making an exhausted research with more data.

Let distribution of travel time on link i be $Dist_i$, estimated based on data sets TT_i^{ind} and TT_i^{corr}. TT_i^{ind} is a set of travel time realizations on link i from vehicles which do not continue to travel on the path and exit the path after link i. TT_i^{corr} is a set of travel time realizations on link i from vehicles traversing to the next link on the path. In this project, we intend to use the Markov Chain Monte Carlo simulation approach. To establish Markov chain transitional probabilities are computed, similar to the studies in (*106*) and (*107*). In the reference (*106*) and (*107*), only TT_i^{corr} and TT_{i+1}^{corr} sets are used to compute the transitional probabilities. However, using partial will understate the true distribution and the correlation. As not all the trajectories are available to begin with, a further subset will reduce the accuracy of estimations.

In this study, the variability in path travel time is based on a new correlation structure formulated to incorporate spatial and temporal parameters. Furthermore, moving away from pre-defined time-based analysis of correlation, travel time correlation is determined among the neighboring links as a function of arrival times at the links.

$$Var\left(\Theta_P(T)\right) = Var\left(\sum_{i=1}^{n} \theta_i(\tau_i)\right) = \sum_{i=1}^{n} Var\left(\theta_i(\tau_i)\right) + 2 \sum_{\substack{1 \le i < j \le n \\ |i-j| \le o}} Cov\left(\theta_i(\tau_i), \theta_j\left(\tau_j\right)\right) \quad (9\text{-}15)$$

$$\tau_j = T + \sum_{i=1}^{j-i} \theta_i(\tau_i)\,, \forall l > 1 \quad (9\text{-}16)$$

where,

$\Theta_P(T)$: The travel time of Path P at time T

o : ($\mathbb{Z}^{\geq}$) The order is the maximum number of neighboring links that are considered to be correlated

$\theta_i(\tau_i)$: Travel time on link i when arrival time is τ_i

τ_i : Estimated arrival time at link i if departure from the first link is at time T

τ_1 : Arrival time at link 1, it is the departure time, i.e., T

The variance, $Var\left(\Theta_P(T)\right)$, in travel time, $\Theta_P(T)$, over path P at departure time T is estimated according to equation (9-6). $Cov\left(\theta_i(\tau_i), \theta_j(\tau_j)\right)$ is the covariance of travel times, $\theta_i(\tau_i)$ and $\theta_j(\tau_j)$, on links i and j when the estimated arrival times at the links are τ_i and τ_j. Arrival times are estimated as per equation (9-7). Arrival at the first link is the departure time, hence, it is equal to T.

9.9. EXPERIMENT RESULTS

The results are produced with the simulated data as described in the section 9.2 above. A path consisting of 14 links was adopted to perform analysis on. Figure 9-9 through Figure 9-11 present temporal profile and other distribution characteristics of travel time on the path and constituting links.

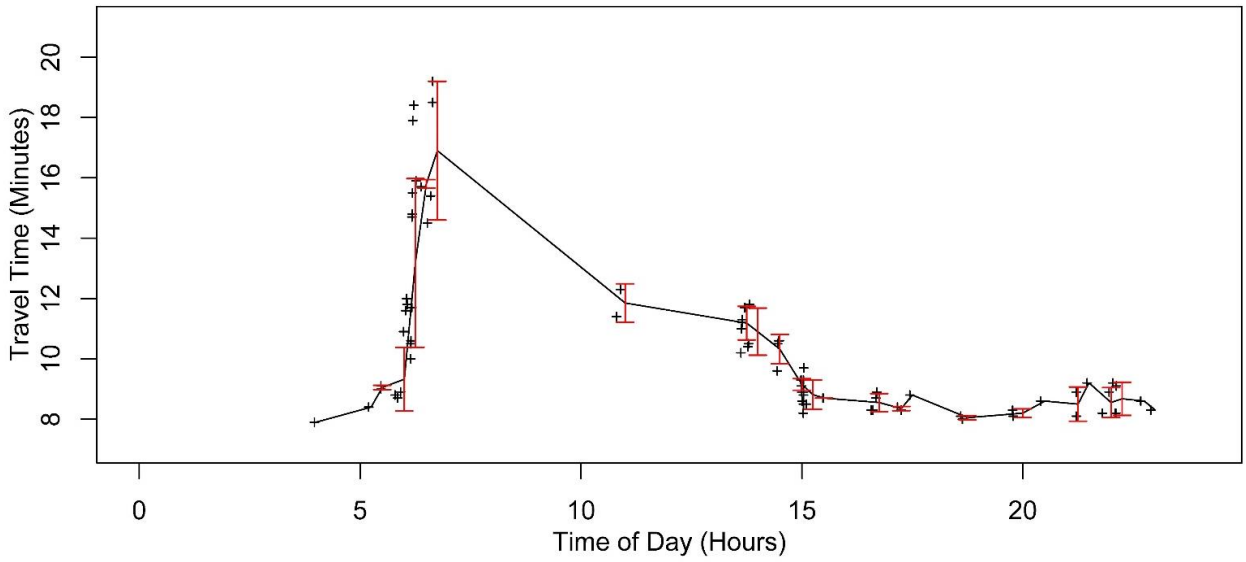


Figure 9-9 Temporal profile of the path travel time distribution.

Path travel times were retrieved from the vehicles that traveled the whole path, and for the link travel times, vehicles which traveled a part of the path were also considered.

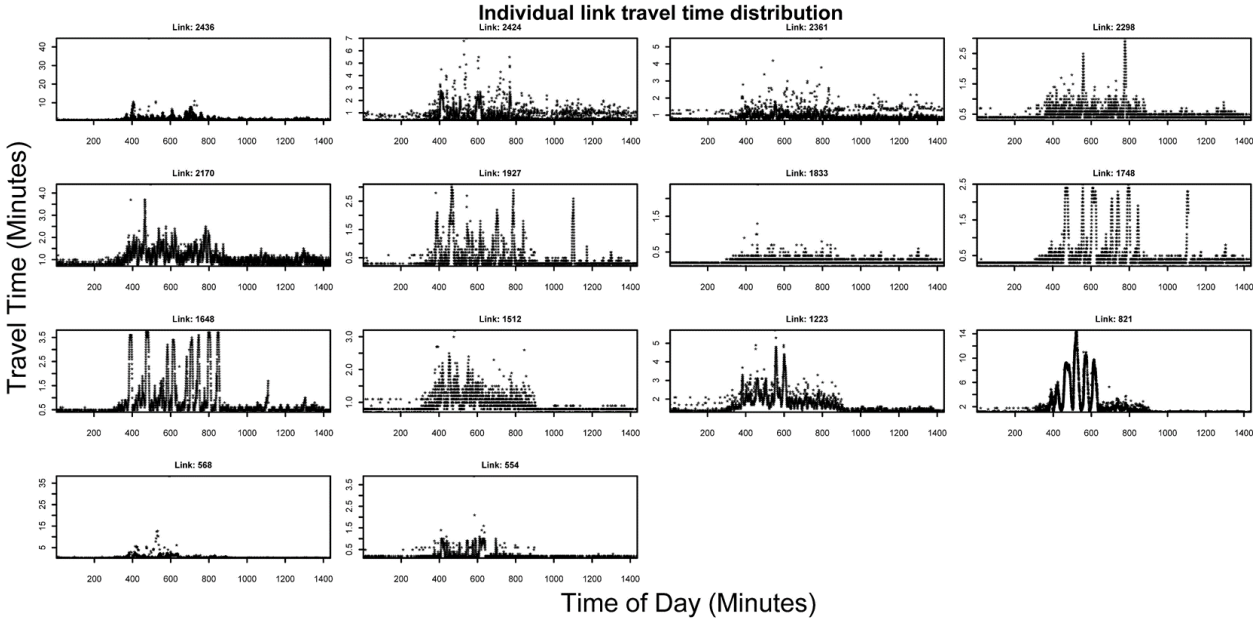


Figure 9-10 Temporal profile of link travel times on the selected path.

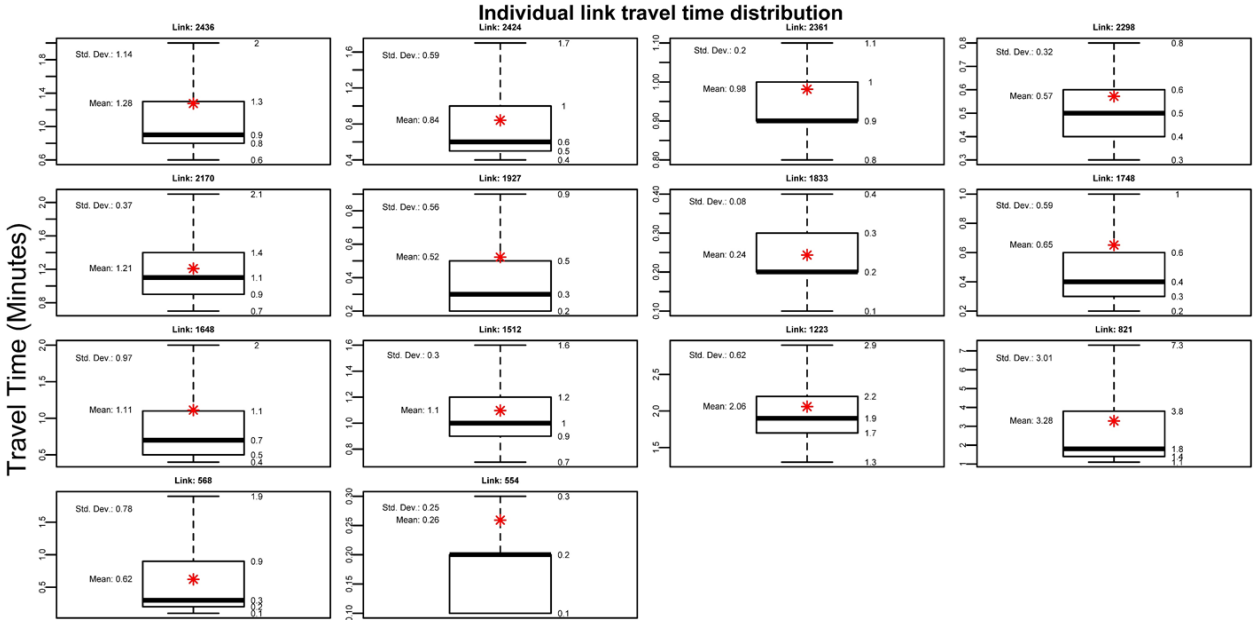


Figure 9-11 Distribution of link travel times on the selected path.

The three techniques mentioned above were implemented to reconstruct path travel times from the corresponding link travel times. Results are presented in Figure 9-12 through Figure 9-14.

9.9.1. Independent Link Travel Time

In this step, link travel times were assumed to be independent, i.e., uncorrelated. This step is based on the following iterative procedure:

- One realized travel time was picked randomly for each of the links. A set of link travel times is constructed.
- Picked link travel times are summed together to estimate a path travel time
- This process is reiterated multiple time to obtain a set of path travel time estimates

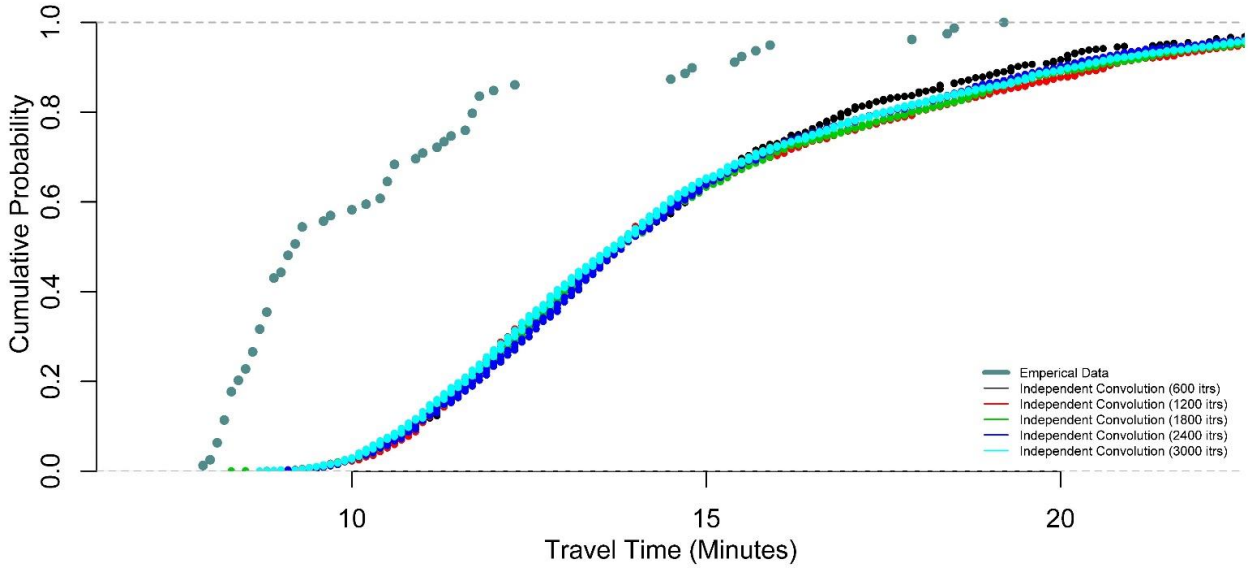


Figure 9-12 Travel time distribution on the path assuming independent link travel times.

Different numbers of iterations were tested to study the effect. There was no major change among different cases of iterations. As it can be observed from the Figure 9-12, assuming independent link travel times results in an inaccurate estimate of path travel time.

9.9.2. Quasi-Time Dependent Link Travel Times

Here, the link travel times were assumed to be quasi-time dependent. The steps involved in this process are as follows:

- Whole time duration is binned in uniform 15-minute intervals.
- A time bin is chosen randomly.
- A realized link travel time for each of the constituting links is picked from the chosen time bin to form a vector.
- Elements of the vector are summed to get a path travel time estimate for the time bin.
- The process is repeated multiple times.

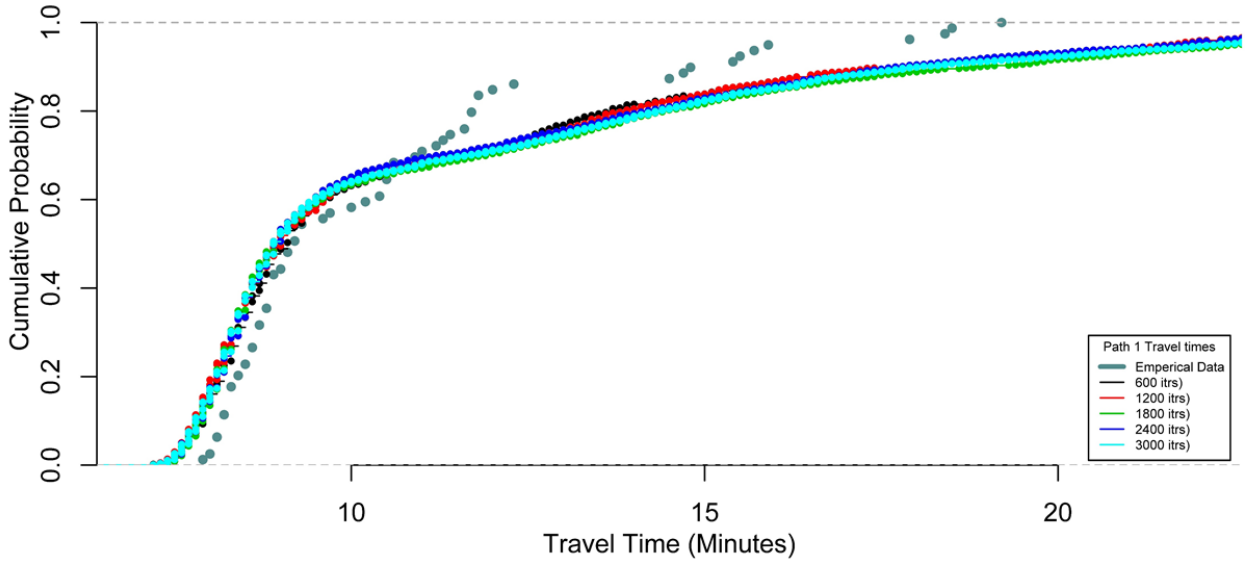


Figure 9-13 Travel time distribution on the path assuming quasi-time dependent link travel times.

Various numbers of iterations were tested to study the effect. With the increase in iterations, estimate gets closer to the empirical distribution relaxing the assumption of independent link travel time's results in the better estimate (Figure 9-13).

9.9.3. Time-Dependent Link Travel Times

This step is based on the previous step of quasi-time dependent link travel times. In the previous step, link travel times from the same (departure time) bin, say the 10:30 AM to 10:45 AM bin, were selected to produce a travel time for the path for that same bin. However, this is not true time-dependence, because the bin for a downstream link, in reality, be different; for example, if the travel time from link 1 to link 5 is more than 15 minutes (technically 7.5minutes if selected a point at random), then one would be in a different time bin for the downstream link. Time-dependence is about recognizing that the arrival time at a downstream node depends on the travel time along link or sub-path to that node. This is likely to be even more of an issue for the shorter time bins (e.g., 2-minute). To incorporate the true time dependency, the time bins were adjusted based on the arrival time at the start node of the link.

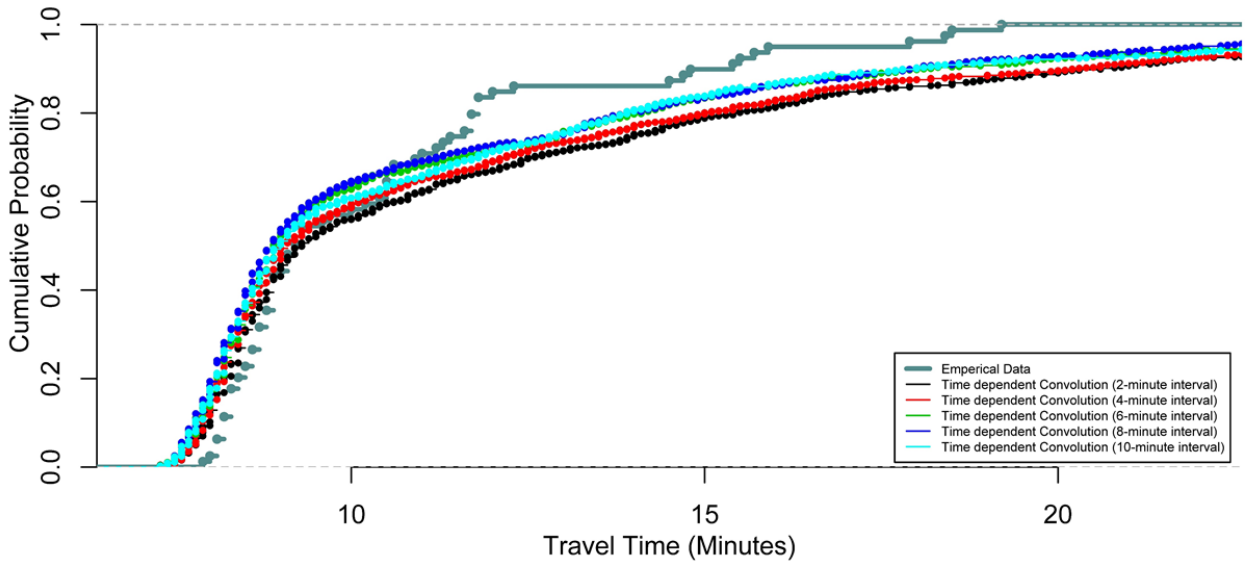


Figure 9-14 Travel time distribution on the path assuming time-dependent link travel times.

With the corrected time bin, the resulting path travel time distribution was estimated better than the previous steps Figure 9-14.

9.9.4. Time-Dependent Link Travel Times

This step is based on the previous and incorporates correlation in link travel times. It can be observed from Figure 9-15 that the estimates are better. In the previous step, the estimated distribution overestimated the travel time. The estimates are very close to the empirical distribution.

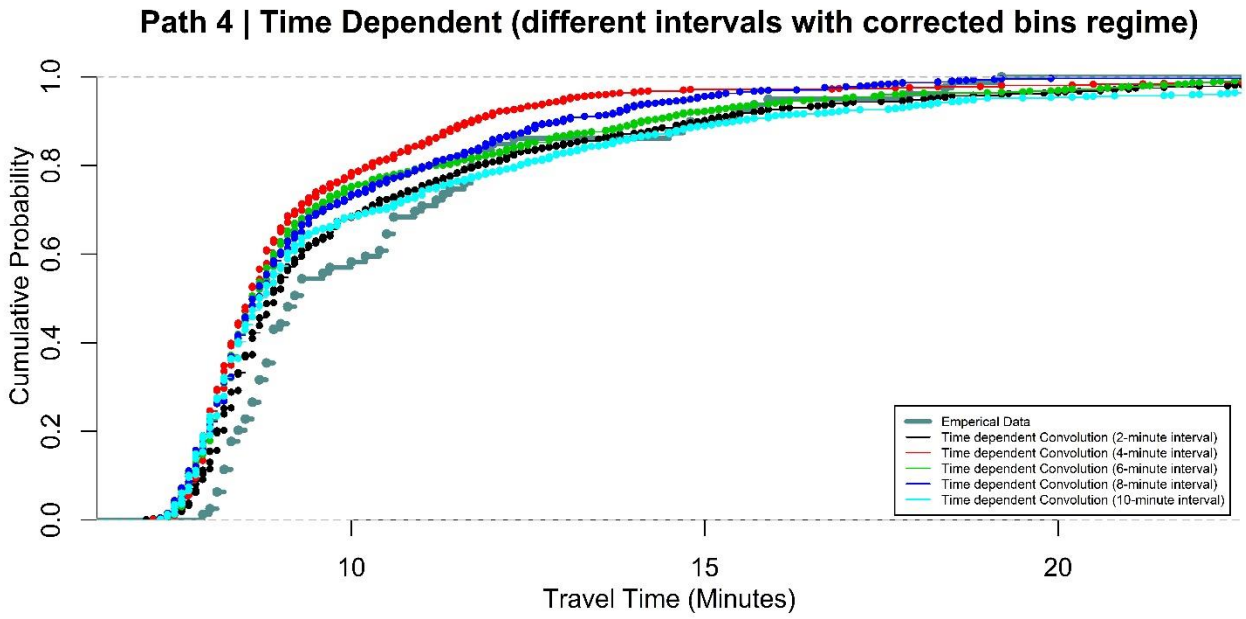


Figure 9-15 Travel time distribution on the path assuming time-dependent link travel times.

9.9.5. The variance of Travel Time

Figure 9-16 presents the estimated variance against the empirical. The empirical and the estimated data were binned into 6-minute intervals. It was found that with an increase in the interval size estimates were off from the empirical data. This is because the when the interval size is larger than the path's travel time, the estimated arrival times at subsequent links fall in the same time bin. This results in wrong sampling and hence the wrong estimate of the variance. Hence, the size of the time bin is critical. Essentially, smaller time bin would be ideal, but with smaller bins, odds of having enough data points gets lowered.

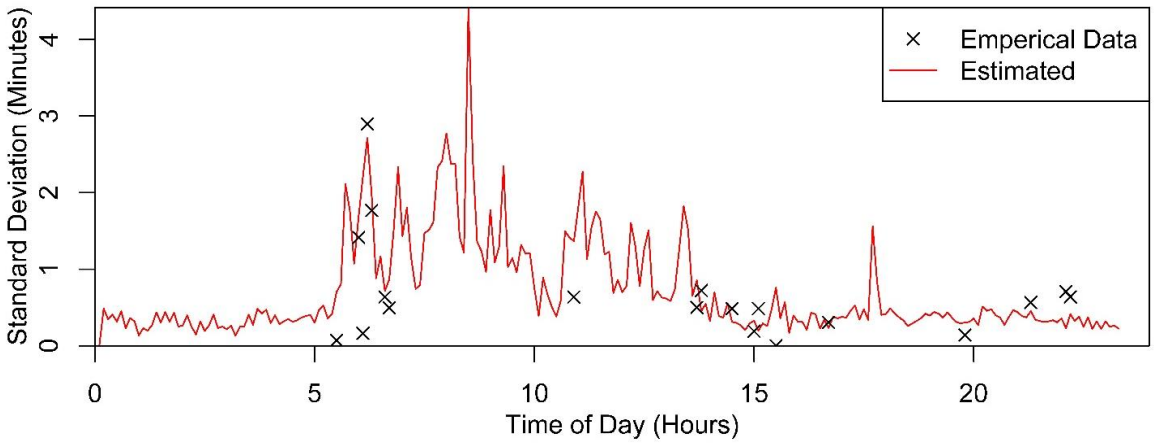


Figure 9-16 Travel time variability.

9.10. Conclusion

In this chapter a travel time distribution along the paths defined by users was estimated through solving a convoluting integral of correlated link travel times. Travel time reliability is an important measure to assess a system's performance. Underlying all measures of travel time reliability is the variability of travel times, experienced along individual links, by various modes, at junctions and intermodal terminals, through the individual trajectories of travelers and goods. The main factors known to affect travel time variation, such as congestion and weather, tend to act on multiple links simultaneously, and their impact tends to linger long after the event itself may have cleared.

In this chapter a solution method was developed to incorporate link travel time covariation. Estimating general covariance patterns is challenging because it requires a much more significant number of observations than estimating the first or second moments of these distributions. Also, obtaining the path-level distributions requires convolving the link-level distributions, a process

that typically does not have closed-form analytical solutions, and for which numerical integration techniques may not always converge.

In this chapter a new solution technique is formulated to convolve link travel times to estimate path travel time. The methodology explicitly incorporated time varying covariance structure. Furthermore, an analytical form of variance of path was devised to correctly capture the spatio-temporal covariance of link travel times.

Results from numerical experiments show that the formulated methods estimate travel times along a route close to the empirical distributions of travel time and its temporal variation.

CHAPTER 10. CONCLUSION AND FUTURE WORK

The primary objective of this dissertation is to explore the potential of CAV technology in formulating 1) facility type-customized active management and control strategies, 2) assessment techniques, and 3) analytical methods with the purpose of enhancing existing transportation network's operational capabilities. Particular attention was dedicated to investigating the opportunities to improve the mobility of vehicles, the efficiency of transportation systems, and environmental standards in a meaningful manner when only a fraction of the traffic stream is connected.

This study models a connected environment to emulate data transmission and explores its potential as, both, a data source and an application platform. Furthermore, connected vehicle-generated traffic data was established as the basis of traffic state characterization and reactive and predictive analytics under different operational conditions. The conducted analysis disaggregates a transportation network in multiple ways to account for various aspects such as demand levels, operational conditions, types of facility and levels of analysis.

To explore the potential of CAVs, the dissertation is divided into three major components. These components address different technical issues of research at different parts of a transportation network. First part evaluates the impact CAVs at freeways and highways via hybrid simulation framework developed in the study. The second component focuses on developing traffic signal control strategies for mixed vehicular traffic stream on arterial roads. The third component develops numerical assessment techniques for planners to understand a network's performance through user-defined paths.

The current state of the transportation system does not have a significant number of deployed CVs, which in turn means there is not an adequate amount of data from such vehicles. To conduct such analysis, simulation-based studies are being conducted. Hence, the studies conducted in the dissertation are simulation based. However, the methodologies developed are independent of any simulation tool and can be applied with any tool with adequate models encoded in them. Calibration of models was not the focus of the study. Drivers' behavior was emulated at the microscopic level with well-calibrated models adopted from the literature.

The first part of the study analyzes at a large network level to assess the impact of connected vehicles on the freeway and network level. To be able to consider facility or network-level effects of connectivity, vehicle communication and interactions are to be captured at the individual vehicle level. Such microscopic analysis would be extremely computationally costly for large-scale networks. Hence, to model and investigate the impact of connectivity on transportation systems and its impending applications, a framework to integrate microscopic aspects of individual vehicle interactions and drivers' behavior within a mesoscopic simulation tool was proposed. Therefore, microsimulation-based traffic stream variables are characterized and utilized to calibrate mesoscopic models under different MPRs of CVs. Mesoscopic models are then used to simulate and analyze facility or network-level impact of connected vehicles. With the integration, we achieve computational economy of mesoscopic level for a large-scale network while capturing driver's behavior at the individual level. Hence, it is a balanced trade-off between accuracy and efficiency of estimation. In the overall analysis of the connected vehicle environment's impact on transportation system state characterization, network-level fundamental diagrams were applied. Observations from the simulated traffic data show that with an increase in MRP of connected

vehicles, the network operates at a lower maximum density and exhibits an increased flow rate for the same density level. Thus, a highly connected environment has the potential to help a congested network recover from flow breakdown and avoid gridlock. Connected vehicles help reduce mean travel time while making the system more reliable. Connectivity can improve a system's performance by increasing throughput and enhance travel time reliability at all demand levels. The analysis also confirms that the linear relationship between distance weighted travel time and distance weighted standard deviation holds for a network and is not affected by either demand level or Market Penetration Rate (MPR) of connected vehicles. Hence, the network appears to retain its inherent properties (signature).

The second part of the study focuses on developing traffic signal control strategies to manage mixed traffic in a connected environment. The study addresses the question whether mixed traffic (automated vehicles, connected vehicles, and regular vehicles) operations could be improved, and if so, how significant such an improvement might be, when only a fraction of vehicles are connected. Accordingly, two types of control strategies were developed: 1) Prediction-based adaptive control strategy, 2) Real-time platoon self-identification control strategy. The current state of practice in the realm of traffic control is based on the fixed sensors which have limited coverage, and the estimation of the traffic state is not accurate, especially queue lengths. This study overcomes such shortcoming in the current state of practice by developing signal controls with connected vehicles as the source of data in addition to the controller's vehicle tracking capability.

Prediction-based adaptive control method formulated in this research optimizes timing plan settings based on anticipated demand. Vehicle trajectories provide the basis to anticipate demand

and optimize timing plans. Also, information on upcoming signal indications is then utilized for vehicles to optimize their trajectories to reduce unnecessary idling and breaking. Connected and automated vehicles communicate with the signals to convey their desired turning movement and accumulated delay along the corridor. This allows for the isolated controller to operate in an intelligent, yet, fully-actuated manner, recognizing the need to coordinate major direction traffic flows, i.e., to enable progression along the corridor, when warranted. Numerical experiments were conducted to evaluate the performance of the strategy. The strategy was compared to vehicle actuated signal control strategy, which is currently the state of art in practice. Admittedly, the effectiveness of the strategy may have been comparably less significant had there been a more advanced control logic to compare against. The pace of advances in the field urges a more advanced solution and future applications need to examine the multifaceted nature of data sources anticipated to be available in the near future.

Real-time platoon self-identification control strategy represents a *platoon-phase scheduling heuristic* that considers clusters of vehicles as critical jobs. The framework devises an advanced, online, signal control logic for mixed traffic environments utilizing the information from CAVs to augment controller/sensor data. A prerequisite of such an approach is the application of the innovative procedure for segmentation of traffic flows based on CV trajectory data. The positive impact of the strategy was demonstrated by investigating the trend of several performance indicators with respect to the base case. To compare the proposed adjusted spatial longitudinal variation clustering technique against a more conventional approach – a critical headway-based platooning - was also examined within the same control logic. The method differs from what transportation research and practice consider platooning. Reported results correspond to two

analysis levels – isolated intersection and corridor. At corridor level, conventional gap-out platoon-based control, unlike the ASLV self-identification control method, fails to consistently achieve superior operational efficiency compared to the vehicle-actuated type of control.

When comparing both signal control strategies developed in this work, the prediction-based adaptive control method provides a marginally greater benefit, with an additional up to 2% decrease in the delay under the cases tested. However, this strategy tends to be computationally heavy. In future work, one can multithread the process where each scenario can be run simultaneously, considerably reducing the required simulation time. Further complex algorithms could also help cut down the simulation and optimization time.

At an isolated intersection level, the oversaturated period proves to be the most challenging to operate satisfactorily. When the traffic stream consists of 100% regular vehicles, both strategies, compared to the vehicle actuated base-case, slightly underperform during this period. However, it should be noted that the method was explicitly devised for mixed environments and that in all other demand and traffic mix cases proposed strategy substantially outperforms the alternative. The strategies are based on the following new features:

- Vehicle-based computation of performance metrics to optimize control parameters. This provides better insight into the prevailing condition resulting in better optimization of the control parameters.
- “Intelligent fully-actuated” controller logic was developed that utilizes high-resolution information of the vehicles’ state.

- Accumulated delay over a facility forms the basis of the objective function with enables flow synchronization through decentralized logic.
- In real-time platoon-based control strategy, platoons were self-identified based on the vehicles prevailing state rather than a conventional threshold headway.

This study was conducted in a simulated environment, and the strategy was found to reduce delay, travel time rate, and queue buildups.

It is evident that a driver is not restricted in taking only one type of facility when driving from origin to destination. Accordingly, in the third part of the study user-defined paths were studied to identify the characteristics associated with those paths. To determine a set of popular route choices a data-driven trajectories clustering method was utilized. Path Fundamental Diagrams (PFDs) are obtained for each of the identified popular paths. A two-dimensional time correlation (TDTC) method was adopted to group the popular paths into a cluster with similar traffic stream variables key relationships. The study puts forward a procedure that can be applied to divide regional networks into sub-networks based on demand patterns' dynamics. Correspondingly, time-of-day control and management strategies can be tailored to suit specific paths' operational characteristics. An online application, can, potentially, be deployed for real-time traffic flow control.

Another fundamental question that arises from the perspective of a given user (i.e., a traveler or good shipping company) pertains to the variability of travel times along specific paths contemplated by the user. The mere addition of expected travel time over the links to obtain a path's travel time is not sufficient. It is essential to capture and understand the correlation among the link travel times for studies that measure reliability in the system. Accordingly, an analytical

framework was designed to adequately compute link travel time correlations with the aim of accurately determining path-based travel time distributions. Travel time distribution along the paths defined by users was estimated through solving a convoluting integral of correlated link travel times. The solution method was developed to incorporated time-varying covariation structure explicitly. Furthermore, an analytical form of the variance of the path was devised to correctly capture the spatiotemporal covariance of link travel times. Results from numerical experiments show that the formulated methods estimate travel times along a route close to the empirical distributions of travel time and its temporal variation.

This dissertation incorporated the emerging technology of CAVs in traditional transportation system practices. Different aspects of the transportation system ranging from simulation techniques to designing control strategies and finally, analytical methods were addressed. In future work, when the real-world data is available from CAVs, recalibration of the models will provide a more realistic magnitude of the impact of the technology. The current work will provide a tangible benchmark. The following tasks are recommended for future investigation:

- Integrate the developed signal controls for arterial roads with freeway and highway hybrid simulation framework. This integration will help study the impact of the CAVs on the whole network.
- Combine different tools into one composite tool. The combination of tools will provide researchers, planners, and practitioners a common ground to test different aspects of the emerging technology.

- With real-world data, recalibrate the driver behavior models to emulate to the reality as closely as possible.
- Continuing the assessment work at path level, develop a user-friendly interface for planners to monitor system's performance in real-time. Further, develop tools to predict an event such as congestion, breakdown, and gridlock.

REFERENCES

[1] Saberi, M., H. Mahmassani, T. Hou, and A. Zockaie. Estimating Network Fundamental Diagram Using Three-Dimensional Vehicle Trajectories: Extending Edie's Definitions of Traffic Flow Variables to Networks. *Transportation Research Record: Journal of the Transportation Research Board,* No. 2422, 2014, pp. 12-20.

[2] Talebpour, A., H. S. Mahmassani, and S. H. Hamdar. Speed Harmonization Evaluation of Effectiveness Under Congested Conditions. *Transportation Research Record*, No. 2391, 2013, pp. 69-79.

[3] Talebpour, A., H. S. Mahmassani, and S. H. Hamdar. Modeling Lane-Changing Behavior in a Connected Environment: A Game Theory Approach. *Transportation Research Procedia,* Vol. 7, 2015, pp. 420-440.

[4] Talebpour, A., and H. S. Mahmassani. Influence of Autonomous and Connected Vehicles on Stability of Traffic Flow.In *Transportation Research Board of National Academies*, Washington DC, 2015.

[5] Wang, M., W. Daamen, S. P. Hoogendoorn, and B. van Arem. Rolling horizon control framework for driver assistance systems. Part II: Cooperative sensing and cooperative control. *Transportation Research Part C: Emerging Technologies,* Vol. 40, No. 0, 2014, pp. 290-311.

[6] Ge, J. I., and G. Orosz. Dynamics of connected vehicle systems with delayed acceleration feedback. *Transportation Research Part C: Emerging Technologies,* Vol. 46, No. 0, 2014, pp. 46-64.

[7] McGurrin, M., M. Vasudevan, and P. Tarnoff. Benefits of dynamic mobility applications : preliminary estimates from the literature.In, Intelligent Transportation Systems Joint Program Office, 2012.

[8] Buisson, C., and C. Ladier. Exploring the Impact of Homogeneity of Traffic Measurements on the Existence of Macroscopic Fundamental Diagrams. *Transportation Research Record: Journal of the Transportation Research Board,* Vol. 2124, 2009, pp. 127-136.

[9] Geroliminis, N., and C. F. Daganzo. Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B: Methodological,* Vol. 42, No. 9, 2008, pp. 759-770.

[10] Mahmassani, H. S., J. C. Williams, and R. Herman. Investigation of Network-level Traffic Flow Relationships: Some Simulation Results. *Transportation Research Record: Journal of the Transportation Research Board,* Vol. 971, 1984, pp. 121-130.

[11] Williams, J. C., H. S. Mahmassani, and R. Herman. Analysis of Traffic Network Flow Relations and Two-Fluid Model Parameter Sensitivity. *Transportation Research Record: Journal of the Transportation Research Board*, 1985, pp. 95-106.

[12] Saberi, M., and H. Mahmassani. Exploring Properties of Networkwide Flow-Density Relations in a Freeway Network. *Transportation Research Record: Journal of the Transportation Research Board,* Vol. 2315, 2012, pp. 153-163.

[13] Saberi, M., and H. S. Mahmassani. Hysteresis and Capacity Drop Phenomena in Freeway Networks: Empirical Characterization and Interpretation. *Transportation Research Record: Journal of the Transportation Research Board,* Vol. 2391, 2013, pp. 44-55.

[14] Smeed, R. J. Road Capacity of City Centers. *Traffic Engineering and Control*, 1966, pp. 455-458.

[15] Thomson, J. M. Speeds and Flows of Traffic in Central London: 2. Speed–Flow Relations. *Traffic Engineering and Control,* Vol. 8, No. 12, 1967, pp. 721–725.

[16] Wardrop, J. G. Journey Speed and Flow in Central Urban Areas. *Traffic Engineering and Control*, Vol. 9, No. 11, 1968, pp. 528–532.

[17] Godfrey, J. W. The Mechanism of a Road Network. *Traffic Engineering and Control*, Vol. 11, No. 7, 1969, pp. 323–327.

[18] Herman, R., and I. Prigogine. A Two-Fluid Approach to Town Traffic. *Science,* Vol. 204, No. 4389, 1979, pp. 148–151.

[19] Chang, M.-F., and R. Herman. Trip Time Versus Stop Time and Fuel Consumption Characteristics in Cities. *Transportation science,* Vol. 15, No. 3, 1981, pp. 183-209.

[20] Daganzo, C. F. Urban gridlock: Macroscopic modeling and mitigation approaches. *Transportation Research Part B: Methodological,* Vol. 41, No. 1, 2007, pp. 49-62.

[21] Ji, Y., W. Daamen, S. Hoogendoorn, S. Hoogendoorn-Lanser, and X. Qian. Investigating the Shape of the Macroscopic Fundamental Diagram Using Simulation Data. *Transportation Research Record: Journal of the Transportation Research Board,* Vol. 2161, 2010, pp. 40-48.

[22] Mazloumian, A., N. Geroliminis, and D. Helbing. The spatial variability of vehicle densities as determinant of urban network capacity. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences,* Vol. 368, No. 1928, 2010, pp. 4627-4647.

[23] Gayah, V. V., X. Gao, and A. S. Nagle. On the impacts of locally adaptive signal control on urban network stability and the Macroscopic Fundamental Diagram. *Transportation Research Part B: Methodological,* Vol. 70, 2014, pp. 255-268.

[24] Nagle, A. S., and V. V. Gayah. A method to estimate the macroscopic fundamental diagram using limited mobile probe data. *Intelligent Transportation Systems - (ITSC), 2013 16th International IEEE Conference on*, 2013, pp. 1987-1992.

[25] Stevanovic, A. *Adaptive traffic control systems: domestic and foreign state of practice*. 2010.

[26] Day, C. M., and D. M. Bullock. Detector-Free Signal Offset Optimization with Limited Connected Vehicle Market Penetration: Proof-of-Concept Study. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2558, 2016, pp. 54-65.

[27] Priemer, C., and B. Friedrich. A Decentralized Adaptive Traffic Signal Control Using V2I Communication Data. *2009 12th International Ieee Conference on Intelligent Transportation Systems (Itsc 2009)*, 2009, pp. 765-770.

[28] Xie, X. F., S. F. Smith, L. Lu, and G. J. Barlow. Schedule-driven intersection control. *Transportation Research Part C-Emerging Technologies,* Vol. 24, 2012, pp. 168-189.

[29] Goodall, N. J., B. L. Smith, and B. Park. Traffic Signal Control with Connected Vehicles. *Transportation Research Record*, No. 2381, 2013, pp. 65-72.

[30] Feng, Y. H., K. L. Head, S. Khoshmagham, and M. Zamanipour. A real-time adaptive signal control in a connected vehicle environment. *Transportation Research Part C-Emerging Technologies,* Vol. 55, 2015, pp. 460-473.

[31] He, Q., K. L. Head, and J. Ding. Multi-modal traffic signal control with priority, signal actuation and coordination. *Transportation Research Part C-Emerging Technologies*, Vol. 46, 2014, pp. 65-82.

[32] ---. PAMSCOD: Platoon-based arterial multi-modal signal control with online data. *Transportation Research Part C-Emerging Technologies*, Vol. 20, No. 1, 2012, pp. 164-184.

[33] Smith, B. L., R. Venkatanarayana, H. Park, N. Goodall, J. Datesh, and C. Skerrit. IntelliDriveSM Traffic Signal Control Algorithms. *University of Virginia*, 2010.

[34] RAMKUMAR, V., P. HYUNGJUN, and S. LEE. Application of IntelliDriveSM to Address Oversaturated Conditions on Arterials.In *TRB 90th Annual Meeting Compendium of Papers DVD. Washington, DC: Transportation Research Board*, 2011. pp. 1-17.

[35] Jing, P., H. Huang, and L. Chen. An Adaptive Traffic Signal Control in a Connected Vehicle Environment: A Systematic Review. *Information*, Vol. 8, No. 3, 2017.

[36] Younes, M. B., and A. Boukerche. Intelligent Traffic Light Controlling Algorithms Using Vehicular Networks. *Ieee Transactions on Vehicular Technology,* Vol. 65, No. 8, 2016, pp. 5887-5899.

[37] Lee, J., B. Park, and I. Yun. Cumulative Travel-Time Responsive Real-Time Intersection Control Algorithm in the Connected Vehicle Environment. *Journal of transportation engineering,* Vol. 139, No. 10, 2013, pp. 1020-1029.

[38] Kari, D., G. Y. Wu, and M. J. Barth. Development of an Agent-Based Online Adaptive Signal Control Strategy Using Connected Vehicle Technology. *2014 Ieee 17th International Conference on Intelligent Transportation Systems (Itsc)*, 2014, pp. 1802-1807.

[39] Ezawa, H., and N. Mukai. Adaptive Traffic Signal Control Based on Vehicle Route Sharing by Wireless Communication. *Knowledge-Based and Intelligent Information and Engineering Systems, Pt Iv,* Vol. 6279, 2010, pp. 280-+.

[40] Chou, L. D., B. T. Deng, D. C. Li, and K. W. Kuo. A Passenger-based Adaptive Traffic Signal Control Mechanism in Intelligent Transportation Systems. *2012 12th International Conference on Its Telecommunications (Itst-2012)*, 2012, pp. 402-405.

[41] Cai, C., Y. Wang, and G. Geers. Vehicle-to-infrastructure communication-based adaptive traffic signal control. *IET Intelligent Transport Systems,* Vol. 7, No. 3, 2013, pp. 351-360.

[42] Gradinescu, V., C. Gorgorin, R. Diaconescu, V. Cristea, and L. Iftode. Adaptive traffic lights using car-to-car communication. *2007 Ieee 65th Vehicular Technology Conference, Vols 1-6*, 2007, pp. 21-25.

[43] Tiaprasert, K., Y. L. Zhang, X. B. Wang, and X. S. Zeng. Queue Length Estimation Using Connected Vehicle Technology for Adaptive Signal Control. *Ieee Transactions on Intelligent Transportation Systems,* Vol. 16, No. 4, 2015, pp. 2129-2140.

[44] Maslekar, N., J. Mouzna, M. Boussedjra, and H. Labiod. CATS: An adaptive traffic signal system based on car-to-car communication. *Journal of Network and Computer Applications,* Vol. 36, No. 5, 2013, pp. 1308-1315.

[45] Chang, H. J., and G. T. Park. A study on traffic signal control at signalized intersections in vehicular ad hoc networks. *Ad Hoc Networks,* Vol. 11, No. 7, 2013, pp. 2115-2124.

[46] Chandan, K., A. M. Seco, and A. B. Silva. Real-time Traffic Signal Control for Isolated Intersection, using Car-following Logic under Connected Vehicle Environment. *World Conference on Transport Research - Wctr 2016,* Vol. 25, 2017, pp. 1613-1628.

[47] Bin Al Islam, S. M. A., and A. Hajbabaie. Distributed coordinated signal timing optimization in connected transportation networks. *Transportation Research Part C-Emerging Technologies,* Vol. 80, 2017, pp. 272-285.

[48] Ahmane, M., A. Abbas-Turki, F. Perronnet, J. Wu, A. El Moudni, J. Buisson, and R. A. Zeo. Modeling and controlling an isolated urban intersection based on cooperative vehicles. *Transportation Research Part C-Emerging Technologies,* Vol. 28, 2013, pp. 44-62.

[49] Shaghaghi, E., M. R. Jabbarpour, R. M. Noor, H. Yeo, and J. J. Jung. Adaptive green traffic signal controlling using vehicular communication. *Frontiers of Information Technology & Electronic Engineering,* Vol. 18, No. 3, 2017, pp. 373-393.

[50] Liu, W. R., G. R. Qin, Y. He, and F. Jiang. Distributed Cooperative Reinforcement Learning-Based Traffic Signal Control That Integrates V2X Networks' Dynamic Clustering. *Ieee Transactions on Vehicular Technology,* Vol. 66, No. 10, 2017, pp. 8667-8681.

[51] Cheng, J. L., W. G. Wu, J. N. Cao, and K. Q. Li. Fuzzy Group-Based Intersection Control via Vehicular Networks for Smart Transportations. *IEEE Transactions on Industrial Informatics,* Vol. 13, No. 2, 2017, pp. 751-758.

[52] Berg, R. Using IntelliDrive SM Connectivity to Improve Mobility and Environmental Preservation at Signalized Intersections. *SAE International Journal of Passenger Cars-Electronic and Electrical Systems,* Vol. 3, No. 2010-01-2317, 2010, pp. 84-89.

[53] Yang, K. D., S. I. Guler, and M. Menendez. Isolated intersection control for various levels of vehicle technology: Conventional, connected, and automated vehicles. *Transportation Research Part C-Emerging Technologies,* Vol. 72, 2016, pp. 109-129.

[54] Guler, S. I., M. Menendez, and L. Meier. Using connected vehicle technology to improve the efficiency of intersections. *Transportation Research Part C-Emerging Technologies,* Vol. 46, 2014, pp. 121-131.

[55] Li, W., and X. G. Ban. Traffic Signal Timing Optimization in Connected Vehicles Environment. *2017 28th Ieee Intelligent Vehicles Symposium (Iv 2017),* 2017, pp. 1330-1335.

[56] Sun, W. L., J. F. Zheng, and H. X. Liu. A capacity maximization scheme for intersection management with automated vehicles. *Papers Selected for the 22nd International Symposium on Transportation and Traffic Theory,* Vol. 23, 2017, pp. 121-136.

[57] Chen, S., and D. Sun. An Improved Adaptive Signal Control Method for Isolated Signalized Intersection Based on Dynamic Programming. *Ieee Intelligent Transportation Systems Magazine,* Vol. 8, No. 4, 2016, pp. 4-14.

[58] Webster, F. V. Traffic signal settings.In, 1958.

[59] Wunderlich, R., C. B. Liu, I. Elhanany, and T. Urbanik. A novel signal-scheduling algorithm with quality-of-service provisioning for an isolated intersection. *Ieee Transactions on Intelligent Transportation Systems,* Vol. 9, No. 3, 2008, pp. 536-547.

[60] Chen, L. W., P. Sharma, and Y. C. Tseng. Dynamic Traffic Control with Fairness and Throughput Optimization Using Vehicular Communications. *Ieee Journal on Selected Areas in Communications,* Vol. 31, No. 9, 2013, pp. 504-512.

[61] Yan, F., M. Dridi, and A. El-Moudni. New vehicle sequencing algorithms with vehicular infrastructure integration for an isolated intersection. *Telecommunication Systems,* Vol. 50, No. 4, 2012, pp. 325-337.

[62] Xie, X. F., G. J. Barlow, S. F. Smith, and Z. B. Rubinstein. Platoon-Based Self-Scheduling for Real-Time Traffic Signal Control. *2011 14th International Ieee Conference on Intelligent Transportation Systems (Itsc)*, 2011, pp. 879-884.

[63] Datesh, J., W. T. Scherer, and B. L. Smith. Using k-means clustering to improve traffic signal efficacy in an IntelliDrive SM environment.In *Integrated and Sustainable Transportation System (FISTS), 2011 IEEE Forum on*, IEEE, 2011. pp. 122-127.

[64] Skabardonis, A., S. Shladover, W.-b. Zhang, L. Zhang, J.-Q. Li, K. Zhou, J. Argote, M. Barth, K. Boriboonsomsin, and H. Xia. Advanced Traffic Signal Control Algorithms.In, 2013.

[65] Beak, B., K. L. Head, and Y. H. Feng. Adaptive Coordination Based on Connected Vehicle Technology. *Transportation Research Record*, No. 2619, 2017, pp. 1-12.

[66] Feng, Y. *Intelligent traffic control in a connected vehicle environment*. The University of Arizona, 2015.

[67] Pandit, K., D. Ghosal, H. M. Zhang, and C. N. Chuah. Adaptive Traffic Signal Control With Vehicular Ad hoc Networks. *Ieee Transactions on Vehicular Technology*, Vol. 62, No. 4, 2013, pp. 1459-1471.

[68] Lin, L. Platoon Identification System in Connected Vehicle Environment.In, 2015.

[69] Lioris, J., R. Pedarsani, F. Y. Tascikaraoglu, and P. Varaiya. Platoons of connected vehicles can double throughput in urban roads. *Transportation Research Part C-Emerging Technologies*, Vol. 77, 2017, pp. 292-305.

[70] Feng, Y. H., M. Zamanipour, K. L. Head, and S. Khoshmagham. Connected Vehicle-Based Adaptive Signal Control and Applications. *Transportation Research Record*, No. 2558, 2016, pp. 11-19.

[71] Jin, Q., G. Y. Wu, K. Boriboonsomsin, and M. Barth. Platoon-Based Multi-Agent Intersection Management for Connected Vehicle. *2013 16th International Ieee Conference on Intelligent Transportation Systems - (Itsc)*, 2013, pp. 1462-1467.

[72] Yang, L., J. M. Hu, Y. J. Li, and X. Pei. Platoon-based Signal Control for Single Intersection under Condition of Intelligent Vehicle Infrastructure Cooperation Systems(IVICS). *2017 4th International Conference on Transportation Information and Safety (Ictis)*, 2017, pp. 325-330.

[73] Li, Z. F., L. Elefteriadou, and S. Ranka. Signal control optimization for automated vehicles at isolated signalized intersections. *Transportation Research Part C-Emerging Technologies*, Vol. 49, 2014, pp. 1-18.

[74] Pourmehrab, M., L. Elefteriadou, S. Ranka, and M. Martin-Gasulla. Optimizing Signalized Intersections Performance under Conventional and Automated Vehicles Traffic. *arXiv preprint arXiv:1707.01748*, 2017.

[75] Le, T., C. Cai, and T. Walsh. Adaptive Signal-Vehicle Cooperative Controlling System. *2011 14th International Ieee Conference on Intelligent Transportation Systems (Itsc)*, 2011, pp. 236-241.

[76] Xu, B., X. G. Ban, Y. G. Bian, J. Q. Wang, and K. Q. Li. V2I based Cooperation between Traffic Signal and Approaching Automated Vehicles. *2017 28th Ieee Intelligent Vehicles Symposium (Iv 2017)*, 2017, pp. 1658-1664.

[77] de Luca, S., R. Di Pace, A. Di Febbraro, and N. Sacco. Transportation Systems with Connected and Non-Connected vehicles: Optimal Traffic Control. *2017 5th Ieee International Conference on Models and Technologies for Intelligent Transportation Systems (Mt-Its)*, 2017, pp. 13-18.

[78] Xiang, J. P., and Z. H. Chen. An adaptive traffic signal coordination optimization method based on vehicle-to-infrastructure communication. *Cluster Computing-the Journal of Networks Software Tools and Applications,* Vol. 19, No. 3, 2016, pp. 1503-1514.

[79] Daganzo, C. F., and N. Geroliminis. An analytical approximation for the macroscopic fundamental diagram of urban traffic. *Transportation Research Part B: Methodological,* Vol. 42, No. 9, 2008, pp. 771-781.

[80] Gonzales, E. J., N. Geroliminis, M. J. Cassidy, and C. F. Daganzo. Allocating city space to multiple transportation modes: A new modeling approach consistent with the physics of transport. *UC Berkeley Center for Future Urban Transport: A Volvo Center of Excellence*, 2008.

[81] Geroliminis, N., and A. Skabardonis. Queue spillovers in city street networks with signal-controlled Intersections.In *89th Transportation Research Board Annual Meeting*, 2010.

[82] Geroliminis, N., and J. Sun. Properties of a well-defined macroscopic fundamental diagram for urban traffic. *Transportation Research Part B: Methodological,* Vol. 45, No. 3, 2011, pp. 605-617.

[83] Mahmassani, H., T. Hou, and M. Saberi. Connecting networkwide travel time reliability and the network fundamental diagram of traffic flow. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2391, 2013, pp. 80-91.

[84] Keyvan-Ekbatani, M., A. Kouvelas, I. Papamichail, and M. Papageorgiou. Exploiting the fundamental diagram of urban networks for feedback-based gating. *Transportation Research Part B: Methodological,* Vol. 46, No. 10, 2012, pp. 1393-1403.

[85] Mittal, A., H. S. Mahmassani, and A. Talebpour. Network Flow Relations and Travel Time Reliability in a Connected Environment. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2622, 2017, pp. 24-37.

[86] Geroliminis, N., and D. M. Levinson. Cordon pricing consistent with the physics of overcrowding. *Transportation and Traffic Theory 2009: Golden Jubilee*, 2009, pp. 219-240.

[87] Hau, T. D. Congestion pricing and road investment. *Road pricing, traffic congestion and the environment: Issues of efficiency and social feasibility*, 1998, pp. 39-78.

[88] Small, K. A., and X. Chu. Hypercongestion. *Journal of Transport Economics and Policy (JTEP)*, Vol. 37, No. 3, 2003, pp. 319-352.

[89] Fosgerau, M., and K. A. Small. Hypercongestion in downtown metropolis. *Journal of Urban Economics*, Vol. 76, 2013, pp. 122-134.

[90] Wu, X., H. X. Liu, and N. Geroliminis. An empirical analysis on the arterial fundamental diagram. *Transportation Research Part B: Methodological*, Vol. 45, No. 1, 2011, pp. 255-266.

[91] Kharrat, A., I. S. Popa, K. Zeitouni, and S. Faiz. Clustering algorithm for network constraint trajectories.In *Headway in Spatial Data Handling*, Springer, 2008. pp. 631-647.

[92] Han, B., L. Liu, and E. Omiecinski. NEAT: road network aware trajectory clustering.In *Distributed Computing Systems (ICDCS), 2012 IEEE 32nd International Conference on*, IEEE, 2012. pp. 142-151.

[93] Palma, A. T., V. Bogorny, B. Kuijpers, and L. O. Alvares. A clustering-based approach for discovering interesting places in trajectories.In *Proceedings of the 2008 ACM symposium on Applied computing*, ACM, 2008. pp. 863-868.

[94] Chen, Z., H. T. Shen, and X. Zhou. Discovering popular routes from trajectories.In *Data Engineering (ICDE), 2011 IEEE 27th International Conference on*, IEEE, 2011. pp. 900-911.

[95] Guo, D., X. Zhu, H. Jin, P. Gao, and C. Andris. Discovering Spatial Patterns in Origin-Destination Mobility Data. *Transactions in GIS*, Vol. 16, No. 3, 2012, pp. 411-429.

[96] Bahbouh, K., and C. Morency. Encapsulating and Visualizing Disaggregated Origin-Destination Desire Lines to Identify Demand Corridors. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2430, 2014, pp. 162-169.

[97] Bahbouh, K., J. R. Wagner, C. Morency, and C. Berdier. TraClus-DL: A Desire Line Clustering Framework to Identify Demand Corridors.In *Transportation Research Board 94th Annual Meeting*, 2015.

[98] Ji, Y., and N. Geroliminis. Spatial and temporal analysis of congestion in urban transportation networks.In *Transportation Research Board Annual Meeting, Washington, DC*, 2011.

[99] Djahel, S., R. Doolan, G.-M. Muntean, and J. Murphy. A communications-oriented perspective on traffic management systems for smart cities: Challenges and innovative approaches. *IEEE Communications Surveys & Tutorials,* Vol. 17, No. 1, 2015, pp. 125-151.

[100] Zheng, F., X. Liu, H. v. Zuylen, J. Li, and C. Lu. Travel Time Reliability for Urban Networks: Modelling and Empirics. *Journal of Advanced Transportation,* Vol. 2017, 2017.

[101] Jenelius, E., and H. N. Koutsopoulos. Travel time estimation for urban road networks using low frequency probe vehicle data. *Transportation Research Part B: Methodological,* Vol. 53, 2013, pp. 64-81.

[102] Kim, J., and H. S. Mahmassani. Spatial and temporal characterization of travel patterns in a traffic network using vehicle trajectories. *Transportation Research Part C: Emerging Technologies,* Vol. 59, 2015, pp. 375-390.

[103] Yildirimoglu, M., and N. Geroliminis. Experienced travel time prediction for congested freeways. *Transportation Research Part B: Methodological,* Vol. 53, 2013, pp. 45-63.

[104] Westgate, B. Vehicle travel time distribution estimation and map-matching via Markov chain monte carlo Methods. 2013.

[105] Westgate, B. S., D. B. Woodard, D. S. Matteson, and S. G. Henderson. Travel time estimation for ambulances using Bayesian data augmentation. *The Annals of Applied Statistics,* 2013, pp. 1139-1161.

[106] Ramezani, M., and N. Geroliminis. On the estimation of arterial route travel time distribution with Markov chains. *Transportation Research Part B: Methodological,* Vol. 46, No. 10, 2012, pp. 1576-1590.

[107] Ma, Z., H. N. Koutsopoulos, L. Ferreira, and M. Mesbah. Estimation of trip travel time distribution using a generalized Markov chain approach. *Transportation Research Part C: Emerging Technologies,* Vol. 74, 2017, pp. 1-21.

[108] Hunter, T., A. Hofleitner, J. Reilly, W. Krichene, J. Thai, A. Kouvelas, P. Abbeel, and A. Bayen. Arriving on time: estimating travel time distributions on large-scale road networks. *arXiv preprint arXiv:1302.6617,* 2013.

[109] Zockaie, A., H. S. Mahmassani, and J. Kim. Network-wide Time-dependent Link Travel Time Distributions with Temporal and Spatial Correlations.In, 2016.

[110] Kwong, K., R. Kavaler, R. Rajagopal, and P. Varaiya. Arterial travel time estimation based on vehicle re-identification using wireless magnetic sensors. *Transportation Research Part C: Emerging Technologies,* Vol. 17, No. 6, 2009, pp. 586-606.

[111] Rahmani, M., E. Jenelius, and H. N. Koutsopoulos. Non-parametric estimation of route travel time distributions from low-frequency floating car data. *Transportation Research Part C: Emerging Technologies,* Vol. 58, 2015, pp. 343-362.

[112] Lu, C. Estimate freeway travel time reliability under recurring and nonrecurring congestion. 2017.

[113] Isukapati, I., G. List, B. Williams, and A. Karr. Synthesizing route travel time distributions from segment travel time distributions. *Transportation Research Record: Journal of the Transportation Research Board,* No. 2396, 2013, pp. 71-81.

[114] Eisele, W. L., B. Naik, and L. R. Rilett. Estimating route travel time reliability from simultaneously collected link and route vehicle probe data and roadway sensor data. *International Journal of Urban Sciences,* Vol. 19, No. 3, 2015, pp. 286-304.

[115] Chen, X., and C. Osorio. Analytical formulation of the trip travel time distribution. *Transportation Research Procedia,* Vol. 3, 2014, pp. 366-373.

[116] Chen, M., G. Yu, P. Chen, and Y. Wang. A copula-based approach for estimating the travel time reliability of urban arterial. *Transportation Research Part C: Emerging Technologies,* Vol. 82, 2017, pp. 1-23.

[117] Mahmassani, H., T. Hou, and J. Dong. Characterizing travel time variability in vehicular traffic networks: deriving a robust relation for reliability analysis. *Transportation Research Record: Journal of the Transportation Research Board,* No. 2315, 2012, pp. 141-152.

[118] Van Lint, J., and H. Van Zuylen. Monitoring and predicting freeway travel time reliability: using width and skew of day-to-day travel time distribution. *Transportation Research Record: Journal of the Transportation Research Board,* No. 1917, 2005, pp. 54-62.

[119] Fei, X., C.-C. Lu, and K. Liu. A bayesian dynamic linear model approach for real-time short-term freeway travel time prediction. *Transportation Research Part C: Emerging Technologies,* Vol. 19, No. 6, 2011, pp. 1306-1318.

[120] Li, R., and G. Rose. Incorporating uncertainty into short-term travel time predictions. *Transportation Research Part C: Emerging Technologies,* Vol. 19, No. 6, 2011, pp. 1006-1018.

[121] Li, C.-S., and M.-C. Chen. Identifying important variables for predicting travel time of freeway with non-recurrent congestion with neural networks. *Neural Computing and Applications,* Vol. 23, No. 6, 2013, pp. 1611-1629.

[122] Innamaa, S. Self-adapting traffic flow status forecasts using clustering. *IET Intelligent Transport Systems,* Vol. 3, No. 1, 2009, pp. 67-76.

[123] Min, W., and L. Wynter. Real-time road traffic prediction with spatio-temporal correlations. *Transportation Research Part C: Emerging Technologies,* Vol. 19, No. 4, 2011, pp. 606-616.

[124] Tsirigotis, L., E. I. Vlahogianni, and M. G. Karlaftis. Does information on weather affect the performance of short-term traffic forecasting models? *International Journal of Intelligent Transportation Systems Research,* Vol. 10, No. 1, 2012, pp. 1-10.

[125] Vlahogianni, E., and M. Karlaftis. Testing and comparing neural network and statistical approaches for predicting transportation time series. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2399, 2013, pp. 9-22.

[126] Castro-Neto, M., Y.-S. Jeong, M.-K. Jeong, and L. D. Han. Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions. *Expert systems with applications,* Vol. 36, No. 3, 2009, pp. 6164-6173.

[127] Giacomini, R., and C. W. Granger. Aggregation of space-time processes. *Journal of econometrics,* Vol. 118, No. 1, 2004, pp. 7-26.

[128] Kamarianakis, Y., and P. Prastacos. Space–time modeling of traffic flow. *Computers & Geosciences,* Vol. 31, No. 2, 2005, pp. 119-133.

[129] Yue, Y., and A. G.-O. Yeh. Spatiotemporal traffic-flow dependency and short-term traffic forecasting. *Environment and planning B: planning and design*, Vol. 35, No. 5, 2008, pp. 762-771.

[130] Cheng, T., J. Haworth, and J. Wang. Spatio-temporal autocorrelation of road network data. *Journal of Geographical Systems,* Vol. 14, No. 4, 2012, pp. 389-413.

[131] Kamarianakis, Y., and P. Prastacos. Forecasting traffic flow conditions in an urban network: Comparison of multivariate and univariate approaches. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1857, 2003, pp. 74-84.

[132] Vlahogianni, E. I., M. G. Karlaftis, and J. C. Golias. Short-term traffic forecasting: Where we are and where we're going. *Transportation Research Part C: Emerging Technologies,* Vol. 43, 2014, pp. 3-19.

[133] Wang, Y., Y. Zheng, and Y. Xue. Travel time estimation of a path using sparse trajectories.In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, 2014. pp. 25-34.

[134] Talebpour, A., H. S. Mahmassani, and F. E. Bustamante. Modeling Driver Behavior in a Connected Environment: Integrated Microscopic Simulation of Traffic and Mobile Wireless Telecommunication Systems. Presented at 95th Annual Meeting of the Transportation Research Board of National Academies, Washington D.C., 2016.

[135] Talebpour, A., H. Mahmassani, and S. Hamdar. Multiregime Sequential Risk-Taking Model of Car-Following Behavior. *Transportation Research Record: Journal of the Transportation Research Board,* Vol. 2260, 2011, pp. 60-66.

[136] Kahneman, D., and A. Tversky. Prospect Theory: An Analysis of Decision under Risk. *Econometrica,* Vol. 47, No. 2, 1979, pp. 263-291.

[137] Hamdar, S., M. Treiber, H. Mahmassani, and A. Kesting. Modeling driver behavior as sequential risk-taking task. *Transportation Research Record: Journal of the Transportation Research Board,* No. 2088, 2008, pp. 208-217.

[138] Hamdar, S. H. Modeling Driver Behavior as a Stochastic Hazard-Based Risk-Taking Process.In *Civil and Environmental Engineering, No. Doctor of Philosophy*, Northwestern University, Evanston, Illinois, 2009. p. 197.

[139] Federal Highway Administration. Next Generation Simulation: US-101 Highway Dataset.In, 2006.

[140] Kesting, A., M. Treiber, and D. Helbing. Enhanced intelligent driver model to access the impact of driving strategies on traffic capacity. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences,* Vol. 368, No. 1928, 2010, pp. 4585-4605.

[141] Treiber, M., A. Hennecke, and D. Helbing. Congested traffic states in empirical observations and microscopic simulations. *Physical review E,* Vol. 62, No. 2, 2000, pp. 1805-1824.

[142] Jayakrishnan, R., H. S. Mahmassani, and T.-Y. Hu. An evaluation tool for advanced traffic information and management systems in urban networks. *Transportation Research Part C: Emerging Technologies,* Vol. 2, No. 3, 1994, pp. 129-147.

[143] Mahmassani, H. S., H. Sbayti, and X. Zhou. Dynasmart-p version 1.0 user's guide. *Maryland Transportation Initiative, College Park, Maryland,* Vol. 137, 2004.

[144] Yelchuru, B., H. S. Mahmassani, I. Zohdy, and R. Kamalanathsharma. Analysis, modeling, and simulation (AMS) testbed development and evaluation to support dynamic mobility applications (DMA) and active transportation and demand management (ATDM) programs -- Chicago calibration report.In, United States. Dept. of Transportation. ITS Joint Program Office, 2016.

[145] Mahmassani, H. S., T. Hou, J. Kim, Y. Chen, Z. Hong, H. Halat, and R. Haas. Implementation of a weather responsive traffic estimation and prediction system (TrEPS) for signal timing at Utah DOT.In, United States. Dept. of Transportation. ITS Joint Program Office, 2014.

[146] Mahmassani, H. S. Autonomous Vehicles and Connected Vehicle Systems: Flow and Operations Considerations. *Transportation science,* Vol. 50, No. 4, 2016, pp. 1140-1162.

[147] Talebpour, A., and H. S. Mahmassani. Influence of connected and autonomous vehicles on traffic flow stability and throughput. *Transportation Research Part C-Emerging Technologies,* Vol. 71, 2016, pp. 143-163.

[148] PTV, A. VISSIM 5.40 user manual. *Karlsruhe, Germany*, 2011.

[149] Stanek, D., R. T. Milam, E. Huang, and Y. A. Wang. Measuring Autonomous Vehicle Impacts on Congested Networks Using Simulation.In, 2018.

[150] Assemblies, T. C. NEMA Standards Publication No. *TS-2, National Electrical Manufacturers Association, Washington, DC*, 1992.

[151] Mahmassani, H. S. Dynamic network traffic assignment and simulation methodology for advanced system management applications. *Networks and spatial economics*, Vol. 1, No. 3-4, 2001, pp. 267-292.

[152] Husch, D., and J. Albeck. Synchro Studio 7 user guide. *Trafficware Ltd., Sugar Land, TX*, Vol. 41, 2006, p. 42.

[153] Shelby, S. Single-intersection evaluation of real-time adaptive traffic signal control algorithms. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1867, 2004, pp. 183-192.

[154] Sims, A. G., and K. W. Dobinson. The Sydney coordinated adaptive traffic (SCAT) system philosophy and benefits. *Ieee Transactions on Vehicular Technology*, Vol. 29, No. 2, 1980, pp. 130-137.

[155] Hunt, P., D. Robertson, R. Bretherton, and R. Winton. SCOOT-a traffic responsive method of coordinating signals.In, 1981.

[156] Gartner, N. H. *OPAC: A demand-responsive strategy for traffic signal control*. 1983.

[157] Mauro, V., and C. Di Taranto. Utopia.In *Control, computers, communications in transportation*, Elsevier, 1990. pp. 245-252.

[158] Mirchandani, P., and L. Head. A real-time traffic signal control system: architecture, algorithms, and analysis. *Transportation Research Part C: Emerging Technologies*, Vol. 9, No. 6, 2001, pp. 415-432.

[159] Luyanda, F., D. Gettman, L. Head, S. Shelby, D. Bullock, and P. Mirchandani. ACS-Lite algorithmic architecture: applying adaptive control system technology to closed-loop traffic signal control systems. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1856, 2003, pp. 175-184.

[160] Brilon, W., and T. Wietholt. Experiences with adaptive signal control in Germany. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2356, 2013, pp. 9-16.

[161] Hong, Z., Y. Chen, H. S. Mahmassani, and S. Xu. Spatial Trajectory Clustering for Potential Route Identification and Participation Analysis for Carpool Commuters. Presented at Transportation Research Board 95th Annual Meeting, 2016.

[162] Hong, Z., S. Xu, and H. S. Mahmassani. Network Topology Aware Moving Object Trajectory Clustering. Presented at Transportation Research Board 2016 Annual Meeting, Washington DC, 2016.

[163] Team, R. C. R language definition. *Vienna, Austria: R foundation for statistical computing*, 2000.

[164] Fourer, R., D. M. Gay, and B. Kernighan. *Ampl*. Boyd & Fraser Danvers, MA, 1993.

[165] Giorgino, T. Computing and visualizing dynamic time warping alignments in R: the dtw package. *Journal of statistical Software*, Vol. 31, No. 7, 2009, pp. 1-24.

[166] Arnott, R. A bathtub model of downtown traffic congestion. *Journal of Urban Economics*, Vol. 76, 2013, pp. 110-121.

[167] Dong, J., H. S. Mahmassani, S. Erdoğan, and C.-C. Lu. State-dependent pricing for real-time freeway management: Anticipatory versus reactive strategies. *Transportation Research Part C: Emerging Technologies*, Vol. 19, No. 4, 2011, pp. 644-657.

[168] Geroliminis, N., A. Srivastava, and P. Michalopoulos. A dynamic-zone-based coordinated ramp-metering algorithm with queue constraints for Minnesota's freeways. *Ieee Transactions on Intelligent Transportation Systems*, Vol. 12, No. 4, 2011, pp. 1576-1586.

[169] Fosgerau, M., and A. Karlström. The value of reliability. *Transportation Research Part B: Methodological*, Vol. 44, No. 1, 2010, pp. 38-49.

[170] Miller-Hooks, E., and H. Mahmassani. Path comparisons for a priori and time-adaptive decisions in stochastic, time-varying networks. *European Journal of Operational Research*, Vol. 146, No. 1, 2003, pp. 67-82.

[171] Miller-Hooks, E. D., and H. S. Mahmassani. Least expected time paths in stochastic, time-varying transportation networks. *Transportation science*, Vol. 34, No. 2, 2000, pp. 198-215.

[172] Nie, Y. M., and X. Wu. Shortest path problem considering on-time arrival probability. *Transportation Research Part B: Methodological*, Vol. 43, No. 6, 2009, pp. 597-613.

[173] Zockaie, A., Y. Nie, and H. Mahmassani. Simulation-based method for finding minimum travel time budget paths in stochastic networks with correlated link times. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2467, 2014, pp. 140-148.

[174] Mahmassani, H. S., J. Kim, Y. Chen, Y. Stogios, A. Brijmohan, and P. Vovsha. *Incorporating Reliability Performance Measures into Operations and Planning Modeling Tools*. Transportation Research Board, 2014.

[175] Mehta, N. B., J. Wu, A. F. Molisch, and J. Zhang. Approximating a sum of random variables with a lognormal. *IEEE Transactions on Wireless Communications,* Vol. 6, No. 7, 2007.

[176] Rakha, H. A., I. El-Shawarby, M. Arafeh, and F. Dion. Estimating path travel-time reliability.In *Intelligent Transportation Systems Conference, 2006. ITSC'06. IEEE*, IEEE, 2006. pp. 236-241.

[177] Sherali, H. D., J. Desai, H. Rakha, and I. El-Shawarby. A discrete optimization approach for locating automatic vehicle identification readers for the provision of roadway travel times.In *82nd Annual Meeting of the Transportation Research Board*, Citeseer, 2002.

APPENDIX I

To estimate the arrival of vehicles at the stop bar and subsequently their exit times, basic speed and acceleration calculations were performed.

With respect to vehicle's arrival time, it is considered to consist of two parts, viz. T_1 and T_2.

First, T_1, is the time vehicles needs to accelerate at the rate Acc to its desired speed, V_{DesSpd}, from its prevailing speed V_f. However, vehicle's distance from the stop bar, $Dist$ may not be sufficient to achieve its desired speed before exiting the intersection. In that case, only the time the vehicle takes while accelerating until arriving at the stop bar is considered.

Second, T_2, is the time vehicles spend on an approach while travelling remainder of their $Dist$ at its desired speed until they arrive at the stop bar.

$$T_1 = \min\left(\frac{V_{DesSpd} - V_f}{Acc}, \frac{-V_f + \sqrt{V_f^2 + 2.Acc.Dist}}{Acc}\right)$$

$$T_2 = \frac{Dist - \left(V_f.T_1 + 0.5 * Acc * T_1^2\right)}{V_{DesSpd}}$$

Hence, the arrival time of the vehicle at the stop bar, T^A is

$$T^A = T_1 + T_2$$

For a platoon P, arrival, exit, and service times are calculated as follows:

Variable Name	***Symbol***	***Formula***
Arrival Time	T_P^A	$\min(T_{veh}^A), veh \in P$
Exit Time	T_P^E	$\max(T_{veh}^A), veh \in P$
Service Time	T_P^{Serve}	$T_P^E - T_P^A$

Intersection clearing (exit) time estimation considers (additional) start-up lost time. A vehicle moving slower than its desired speed is assumed to accelerate until it achieves its desired speed and then continues to travel at its desired speed. For the stopped vehicles start-up lost time until saturation headway is achieved is also incorporated into computation of intersection clearing (exit) time estimate.

Figure 0-1 Calculating the arrival/exit time of vehicles

For vehicles 1 through 6 in Figure 0-1, which are stopped, additional time is introduced when estimating their exit times. Start-up lost time and discharge headway is determined based on Figure 0-2. For instance, for vehicle 5, start-up lost time based on its position in the queue (cumulative headway) calculation is 9.0 (3.2 + 3.0 + 2.8) seconds plus time needed to clear the intersection cross-section once it starts moving.

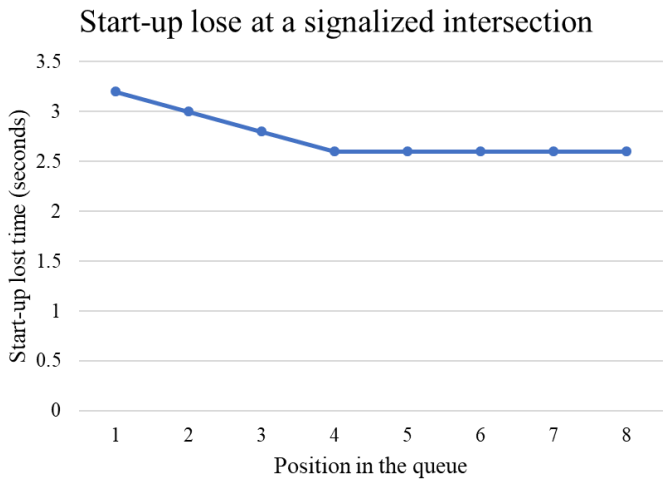


Figure 0-2 Start-up lost time