3-26-2021

Integrated Approach for Diversion Route Performance Management during Incidents

Rajib Chandra Saha
Florida International University, rsaha005@fiu.edu

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

INTEGRATED APPROACH FOR DIVERSION ROUTE PERFORMANCE

MANAGEMENT DURING INCIDENTS

A dissertation submitted in partial fulfillment of

the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

CIVIL ENGINEERING

by

Rajib Chandra Saha

2021

To: Dean John Volakis
College of Engineering and Computing

This dissertation, written by Rajib Chandra Saha, and entitled Integrated Approach for Diversion Route Performance Management during Incidents, having been approved in respect to style and intellectual content, is referred to you for judgment. We have read this dissertation and recommend that it be approved.

Albert Gan

Xia Jin

Priyanka Alluri

Wensong Wu

Yan Xiao

Mohammed Hadi, Major Professor

Date of Defense: March 26, 2021

The dissertation of Rajib Chandra Saha is approved.

Dean John Volakis
College of Engineering and Computing

Andrés G. Gil
Vice President for Research and Economic Development
and Dean of the University Graduate School

Florida International University, 2021

DEDICATION

I dedicate this dissertation to my parents, Dipti Rani Saha and Santi Ranjan Saha, and my wife, Doli Rani Saha, for their unconditional love and support.

ACKNOWLEDGMENT

I pay my deepest homage to Lord Shri Krishna for showering the blessing on me to complete this dissertation.

I want to take this opportunity to acknowledge all the people who supported me in this journey. First and foremost, I would like to express my sincere gratitude to my advisor and mentor, Dr. Mohammed Hadi, for his continuous guidance, kindness, support, and directive suggestions at every stage during my study at Florida International University. I am grateful to him for accepting me into this program, giving me the opportunity to work on exciting projects and grow intellectually.

I would also like to thank all my committee members, Dr. Albert Gan, Dr. Xia Jin, Dr. Priyanka Alluri, Dr. Yan Xiao, and Dr. Wensong Wu, for their guidance in completing my dissertation. I sincerely appreciate spending their valuable time to improve my research.

I acknowledge the Florida Department of Transportation (FDOT) for providing financial support for this research.

All my colleagues at the Lehman Center for Transpiration Research (LCTR) are an integral part of this journey. I owe a big thank you to them for all the good memories throughout this journey.

Finally, I would like to thank my parents Dipti Rani Saha, Santi Ranjan Saha, my parents-in-law Ganga Rani Saha, Bipul Chandra Saha, and my beloved wife Doli Rani Saha, for their endless support, encouragement, and love.

ABSTRACT OF THE DISSERTATION

INTEGRATED APPROACH FOR DIVERSION ROUTE PERFORMANCE MANAGEMENT DURING INCIDENTS

by

Rajib Chandra Saha

Florida International University, 2021

Miami, Florida

Professor Mohammed Hadi, Major Professor

Non-recurrent congestion is one of the critical sources of congestion on the highway. In particular, traffic incidents create congestion in unexpected times and places that travelers do not prepare for. During incidents on freeways, route diversion has been proven to be a useful tactic to mitigate non-recurrent congestion. However, the capacity constraints created by the signals on the alternative routes put limits on the diversion process since the typical time-of-day signal control cannot handle the sudden increase in the traffic on the arterials due to diversion. Thus, there is a need for proactive strategies for the management of the diversion routes performance and for coordinated freeway and arterial (CFA) operation during incidents on the freeway. Proactive strategies provide better opportunities for both the agency and the traveler to make and implement decisions to improve performance.

This dissertation develops a methodology for the performance management of diversion routes through integrating freeway and arterials operation during incidents on the freeway. The methodology includes the identification of potential diversion routes for freeway incidents and the generation and implementation of special signal plans under different incident and traffic conditions. The study utilizes machine learning, data analytics, multi-resolution modeling, and multi-objective optimization for this purpose. A data analytic approach based on the long short-term memory (LSTM) deep neural network method is used to predict the utilized alternative routes dynamically using incident attributes and traffic status on the freeway and travel time on both the freeway and alternative routes during the incident. Then, a combination of clustering analysis, multi-resolution modeling (MRM), and multi-objective optimization techniques are used to develop and activate special signal plans on the identified alternative routes. The developed methods use data from different sources, including connected vehicle (CV) data and high-resolution controller (HRC) data for congestion patterns identification at the critical intersections on the alternative routes and signal plans generation. The results indicate that implementing signal timing plans to better accommodate the diverted traffic can improve the performance of the diverted traffic without significantly deteriorating other movements' performance at the intersection. The findings show the importance of using data from emerging sources in developing plans to improve the performance of the diversion routes and ensure CFA operation with higher effectiveness.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS AND ACRONYMS

AADT	Annual Average Daily Traffic
AAOGR	Average Arrival on Green
ABC	Ant-Based Control
ACA	Ant Colony Algorithm
AMD	Age-related Macular Degeneration
AOG	Arrival on Green
APD	Absolute Percentage Deviation
ATCS	Adaptive Traffic Control System
ATDM	Active Transportation and Demand Management
ATIS	Advanced Traveler Information System
ATMS	Advanced Traffic Management System
ATM	Active Traffic Management
AVL	Automatic Vehicle Location
BSM	Basic Safety Message
CAF	Capacity Adjustment Factors
CCTV	Closed Circuit Television Camera
CFA	Coordinated Freeway and Arterial
CFR	Code of Federal Regulations
CNN	Convolution of Neural Network
COM	Component Object Model
CPU	Central Processing Unit
CV	Connected Vehicle

DBI	Davies-Bouldin Index
DBN	Deep Belief Network
DEC	Deep Embedded Clustering
DMS	Dynamic Message Sign
DNN	Deep Neural Network
DOT	Department of Transportation
DSS	Decision Support System
DTA	Dynamic Traffic Assignment
EKF	Extended Kalman Filter
EM	Expectation Maximization
FCC	Federal Communications Commission
FDOT	Florida Department of Transportation
FHWA	Federal Highway Administration
FLC	Fuzzy-Logic Controller
FTP	File Transfer Protocol
GA	Genetic Algorithm
GIS	Geographic Information System
GPS	Global Positioning System
GOR	Green Occupancy Ratio
HAR	Highway Advisory Radio
HCM	Highway Capacity Manual
HRC	High-Resolution Controller
ICM	Integrated Corridor Management

INDOT	Indiana Department of Transportation
ITIS	Intelligent Traveler Information Systems
ITS	Intelligent Transportation System
KNN	K-Nearest Neighbor
LSTM	Long Short-Term Memory
LTEE	Link Travel Time Estimation Error
MAD	Mean Absolute Deviation
MAE	Mean Absolute Error
MAPD	Mean Absolute Percentage Deviation
MAPE	Mean Absolute Percentage Error
MLP	Multilayer Perceptron
MP	Market Penetration
MRM	Multi-Resolution Modeling
MUTCD	Manual on Uniform Traffic Control Devices
NDP	Network Design Problem
NEMA	National Electrical Manufacturers Association
NSGA-II	Non-dominated Sorting Genetic Algorithm-II
OBU	On-Board Unit
OCT	Optical Coherence Tomography
OD	Origin-Destination
ODME	Origin-Destination Matrix Estimation
OSM	Open Street Maps
PC	Principal Component

PCA	Principal Component Analysis
PCD	Purdue Coordination Diagram
PSO	Particle Swarm Optimization
RBM	Restricted Boltzmann Machines
RITIS	Regional Integrated Transportation Information System
RMSE	Root Mean Squared Error
RMSEP	Root Mean Squared Error Percentage
RNN	Recurrent Neural Network
ROR	Red Occupancy Ratio
RSE	Road Side Equipment
RTMS	Remote Transportation Microwave Sensor
RTSMIP	Real-Time System Management Information Program
SA	Simulated Annealing
SBA	Simulation Based Assignment
SERPM	Southeast Florida Regional Planning Model
SGD	Stochastic Gradient Descent
SNE	Stochastic Neighbor Embedding
SSE	Sum of Square Error
SSNN	State Space Neural Network
TAT	Traffic Analysis Toolbox
TDNN	Time Delayed Neural Network
TIM	Traffic Incident Management
TOD	Time of Day

TIS	Traveler Information System
TMC	Traffic Management Center
TTR	Travel Time Reliability
TSMO	Transportation Systems Management and Operations
t-SNE	t-Distributed Stochastic Neighbor Embedding
USDOT	United States Department of Transportation
V2I	Vehicle to Infrastructure
V2V	Vehicle to Vehicle
XGB	eXtreme Gradient Boosting

CHAPTER I

INTRODUCTION

1.1 Background

Non-recurrent congestion is one of the critical sources of congestion on the highway. Traffic incidents create congestion in unexpected times and places that travelers do not prepare for. A national statistic indicated that around 25% of total congestion occurs due to traffic incidents (FHWA, 2012). Incidents reduce the capacity of the highway (HCM, 2016) by creating bottlenecks that cause additional delays to traffic. For example, a lane blockage for a minute was estimated to result in 4-10 minutes of travel delay (FHWA, 2012). Moreover, the traffic incidents cause secondary crashes (FHWA, 2009), increase fuel consumption, and contribute to air pollution (Yazici et al., 2018).

An incident on a route has the potential to create havoc on the whole corridor or system, including the incident impacted route. In order to efficiently manage the incident and restoring traffic conditions to normal, traffic incident management (TIM) plays a crucial role. Reduction in congestion through better management of incidents has numerous benefits, including reducing travel times, smoothing the traffic flow, increasing average fuel economy, shortening the rush hour period, and reducing vehicle queuing (Mazzenga and Demetsky, 2009).

TIM is the process of managing multi-agency activities in response to highway traffic disruptions. Efficient and coordinated management of incidents is critical to reducing their adverse impacts on safety, traffic operations, and the local economy (FHWA, 2000). TIM includes a variety of incident management strategies to mitigate the effects of the incident on the network. These strategies involve real-time decision making under uncertain and rapidly changing conditions. Transportation agencies utilize cutting-edge intelligent transportation systems (ITS) tools and real-time traffic data in decision making. Some of the strategies that agencies currently apply are Advanced Traveler Information System (ATIS) and Advanced Traffic Management Systems (ATMS), with Active Traffic Management (ATM) and so on.

The ATIS provides travelers with travel time, routing, and other trip information or guidance, which helps the traveler make an informed decision about route choice or departure decision. ATMS offers a significant opportunity to reduce the impact of the incident by reducing incident and lane blockage duration, coordinating incident response activities, and managing route diversion during incidents. ATM dynamically manages the congestion following a continuous process from collecting and analyzing traffic data to selecting and implementing the best strategies. An ideal system comprising these strategies uses prevalent traffic data to predict the near future and suggests a control plan based on the prediction to maximize the effectiveness and efficiency of the network (Kurzhanskiy and Varaiya, 2010). Although incident management by itself is a successful tool to reduce congestion, combining it with traffic signal control strategies is expected to improve the performance significantly (Texas Transportation Institute, 2003).

Traffic incident management coupled with a sound understanding of the incident characteristics provides agencies a vantage point to manage the traffic at the corridor level. Moreover, understanding the nexus between freeways and arterial streets operation and signal control during the incident facilitates the efficient operation of existing facilities with improved performance throughout the corridor.

1.2 Problem Statement

Incident attributes such as incident duration, number of blocked lanes, incident severity, and incident location on the freeway have a considerable effect on vehicles' shift to alternative routes. Alternative routes can include other freeways, major parallel arterials, and frontage roads connected with the impacted freeway through arterials, ramps, etc. These connectors facilitate the detour operation of the vehicle during the incident.

During major incidents, route diversion has been proven to be a useful tactic. The provision of real-time information has the potential to trigger or increase the diversion under incident conditions (Mahmassani, 2001; Ben-Akiva et al., 2002) and consequently has the potential to reduce the overall travel time in the entire network. The success of this technique depends on the proportion of the drivers' compliance with the diversion and the resulting impacts on the alternative routes. A previous study noted that the best network performance occurs with the driver compliance rates between 50 and 60% (Dia and Cottman, 2006). However, this number seems to be high and is expected to be influenced by the conditions on the freeway and the alternative routes. Developing alternate routing plans and strategies for a coordinated operation have become increasingly vital components to realize the diversion's benefits.

One of the crucial factors for effective coordinated strategies is estimating the diverted traffic due to the incident. Based on the stated preference surveys, the estimated diversion is somewhat between 40% and 70% of freeway traffic exiting the freeway ahead of an incident location (Barfield et al., 1989; Benson, 1996; Khattak et al., 1993; Al-Deek et al., 2009). However, when predicting utilizing field detector data, the estimated percentages were found to be between 5% and 25% based on the incident characteristics

(Hadi et al., 2013; Tariq et al., 2019). The resulting congestion on the alternative routes discourages drivers from diverting, resulting in lower percentages of diversion (Khattak et al., 1992). A recent study of diversion based on detector data infers that the capacity of the signals constrains the diversion at the off-ramps during the periods. The study indicates the need for special signal control plans during incidents to increase the capacity of the off-ramps and adjacent signals leading to the primary parallel routes (Tariq et al., 2019).

Therefore, a CFA operation is necessary to accommodate the additional traffic on the alternative routes and concurrently manage the traffic jointly on the freeway and arterial with improved mobility, reliability, and safety. Such coordination should utilize effective and timely signal timing plans to accommodate the diversion. The wrong identification of the incident status and predictions of impacts may lead to switching green to non-deserving movements, which, in turn, causes unnecessary delays to other movements. At intersections with high traffic demands, the consequence of the wrong detection of incident status could be severe (Ahmed and Hawas, 2015). Implementing the wrong response during the incident could also worsen congestion on the directly impacted freeway and its surrounding highway network (Wirtz et al., 2005).

Proactive strategies are more effective in managing the impact of diversion compared to reactive strategies. Proactive strategies will give better opportunities for both the agency and the traveler to make and implement more effective decisions. The Active Transportation and Demand Management (ATDM) approach of FHWA advocates proactive strategies for preventing or delaying breakdown conditions, improving safety, promoting sustainable modes, reducing emissions, or maximizing system efficiency (FHWA, 2021).

The formulation and implementation of proactive strategies require supportive real-time data in high resolution. The existing data collection technologies such as Bluetooth readers, Wi-Fi, INRIX, and HERE act as good sources. Besides, Connected Vehicle (CV) serves as an additional real-time data source. The high-resolution controller (HRC) data has the potential to produce better performance measures of the traffic control system. These new data sources undoubtedly provide a new opportunity for proactive strategies.

Inspired by the availability of a plethora of real-time and archived data and modeling tools, this research focuses on the support of coordinated, proactive strategies using data analytics, predictive methods, and modeling, consequently estimating the benefits of implementing the new management plans. This research provides a feasible framework for coordinated operation of the freeway and arterial during the incident using existing data sources.

1.3 Research Objectives

The primary goal of this research effort is to develop a method for the management of diversion route performance during incidents utilizing the abundance of data from multiple sources through harnessing the power of machine learning combined with modeling. The models and methods developed in the study can be incorporated as part of decision support systems (DSS) to manage the traffic proactively during incidents on the freeway. In this regard, the main objectives of this research are listed below.

- Develop a method for predicting the critical routes utilized by motorists during incidents and the associated traffic and incident conditions under which the diversion occurs due to the incident on the freeway.

- Develop a method for generating special signal timing plans and estimate the benefits of activating those plans to mitigate the deterioration in the performance of the movements of critical route intersections due to diversion.
- Estimate the benefits of utilizing new data sources (i.e., CV and HRC) as part of the methodology to support traffic management in the network.

This research is conducted recognizing that the methodologies currently available to plan, design, and implement detour operation during incidents on the freeway are static and do not consider many influencing dynamic variables such as incidents attributes, traffic attributes, time of day, and existing control. On the other hand, currently considered DSS for coordinated operation of freeway and arterials during incidents are based on the online simulation that many agencies do not have resources or are not comfortable with. The research considers that most of the intersections on arterial streets are operated under time-of-day (TOD) control. The TOD plans are developed based on what are considered normal day conditions and thus cannot handle the sudden demand surges due to diversion. The objectives of this study are set to resolve these issues and to provide a dynamic, proactive, and easily implementable DSS for CFA operation during the incidents in the freeway. Moreover, the study provides a new perspective of the usage of data obtained from various imminent technologies such as CV and HRC for traffic management during incidents.

1.4 Research Organization

This dissertation includes a total of six chapters as follows. Chapter II provides a comprehensive literature review related to this research. Chapter III presents the methodology used to achieve the stated objectives. The data preparation process is described in Chapter IV, and model application and results are discussed in Chapter V. The last chapter summarizes the findings and provides recommendations.

CHAPTER II

LITERATURE REVIEW

The literature review presented in this chapter covers three main areas; the first section covers the impact of incidents and incident management, the second section reviews the modeling of incidents and traffic signal control during incidents. The last section addresses the use of CV and HRC data for traffic management. Gaps in the existing studies and methods for the analyses are identified in this chapter through detailed literature reviews.

2.1 Impacts of Incident

An incident is a stochastic event that creates not only personal losses but also operational hazards. Incidents obstruct the normal traffic flow by blocking the shoulder or lane(s) and thus reducing the capacity. The resulting impacts may not be limited to the location of the incident but can have a substantial influence over the entire corridor that includes the alternative routes. Traffic incidents are significant sources of delay, system unreliability, and efficiency in the corridor.

Lane blockage due to the incident produces a temporary bottleneck by reducing the capacity of the impacted section of the road. This reduction in capacity is not linearly related to the number of blocked lanes. For example, with the blockage of one lane out of three travel lanes, the capacity reduction is not 33%. A study by Smith et al. (2003) on the reduction in freeway capacity due to incidents found 63% capacity reduction for one out of three lanes blocked and 77% reduction for two out of three lanes blocked. The highway capacity manual (HCM, 2016) recommended specific capacity reduction factors due to the lane blockages. Table 2-1 shows the capacity adjustment factors (CAFs) for different levels of lane lockage considering the total number of lanes in the facility. According to HCM 2016, for one lane blockage out of three lanes, the CAF is 0.74, and for two lanes, blockage out of three lanes is 0.52.

Table 2-1: CAFs by Incident Type and Number of Directional Lanes on the Facility

Directional Lanes	No Incident	Shoulder Closed	One Lane Closed	Two Lanes Closed	Three Lanes Closed	Four Lanes Closed
2	1	0.81	0.71	N/A	N/A	N/A
3	1	0.83	0.74	0.51	N/A	N/A
4	1	0.85	0.77	0.5	0.52	N/A
5	1	0.87	0.81	0.67	0.5	0.5
6	1	0.89	0.85	0.75	0.52	0.52
7	1	0.91	0.88	0.8	0.63	0.63
8	1	0.93	0.89	0.84	0.66	0.66

A study using real-world incident scenarios on four-lane freeways in Maryland and Northern Virginia found that the available capacity reduces to 70% of the original capacity in the cases of one-lane closure. For two lanes closure and three lanes closure scenarios, the capacity was reduced to 36% and 17% of the original capacity, respectively. The study analyzed a total of fifty-one incidents from Maryland and Northern Virginia and suggested that changes in driver behaviors and distractions are contributing factors besides the physical blockage of the lanes to capacity reduction (Masghati-Amoli et al., 2015).

Another empirical study using real-world data found that the maximum throughput of the incident location is roughly 50% lower than the flow that could be obtained on the same number of lanes without an incident. In addition to the blockage of the lane, the study indicated two behavioral reasons for this reduction: rubbernecking to watch the crash and lowering the speed to respect the workers' presence. The study also found a high average discharge headway (3 seconds) at the incident location compare to typical headway (1.6 seconds) (Knoop et al., 2008).

The reduction in capacity can result in delays to the users and produce queues on the roadway. The increase in delay continues even after the elapsed time of the incident (Pan et al., 2013). The delay varies by the categories of the incidents (Wang et al., 2008). For major incidents, with durations longer than 60 minutes, one incident can cause 2,500-5,000 veh-hours of delay (Martin et al., 2011). An injury collision typically results in a very long delay on freeways. In this case, the delay can be 274% of the delay caused by a non-injury collision and more than 19 times the delay from a disabled vehicle incident (Wang et al., 2008).

Travel time reliability (TTR) is frequently used to estimate the performance of the freeway, in addition to the average travel time and delay. Intuitively, the incident has an enormous impact on travel time reliability. To quantify the impact, Wright et al. (2015) did an empirical analysis of the travel time reliability on the freeway considering the impacts of the incidents. The travel time variability, buffer index, and the probability of freeway segment traffic breakdown were estimated after the incident and compared with those during normal conditions. Single lane, multilane, and shoulder blockages were considered for the impact analysis. The study indicated that all types of incidents impact travel time reliability, although multiple lane blockage incidents have the highest impacts on travel time reliability.

Another study by Hojati et al. (2016) determined that the TTR is impacted by incident attributes such as the duration, the ratio of flow during the incident to recurrent flow, multiple vehicles involved, and more than one lane blocked.

The incidents on the freeway are also responsible for the occurrence of secondary incidents. A study conducted in Minnesota showed that thirteen percent of all peak-hour crashes are the result of a previous incident (Martin et al., 2011). Several factors, such as primary incident type, primary incident lane-blockage duration, time of day, and weather, have significant effects on the likelihood of a secondary incident (Zhan et al., 2009). Secondary incidents are more likely to occur if the primary incident lasts long; at the same time, the duration of primary incidents is expected to be longer if secondary incidents occur (Khattak et al., 2011). Sometimes primary incident causes multiple secondary incidents. Other than safety impacts, the secondary incidents cause additional delays to the traffic. A study by Khattak et al. (2011) estimated a further 3% delay due to the secondary incident in the Hampton Roads area, Virginia.

2.2 Traffic Incident Management Operation

Since the incident deteriorates the performance of the roadway, multiple agencies work jointly to manage the incident and curb the effects of the incident from spreading over the entire corridor. Managing the incident involves a series of activities carried out by different agencies and organizations. These activities may not be performed sequentially. The following are the detailed process of incident management described in the Freeway Management and Operations Handbook (Neudorff et al., 2003).

Detection: An incident comes to the attention of the agency or agencies responsible for maintaining traffic flow and safe operations on the facility.

Verification: Incident is verified along with other relevant information such as exact location, type, etc. A proper initial response team is dispatched based on the incident.

Motorist Information: Various incident-related information is disseminated via available media as soon as possible to the affected motorist to make them aware of the situation.

Response: Incident response includes dispatching the appropriate personnel and equipment and activating the appropriate communication links and motorist information media as soon as there is reasonable certainty that an incident is present.

Site Management: Site management involves effectively coordinating and managing on-scene resources for ensuring the safety of response personnel, incident victims, and other motorists.

Traffic Management: Traffic management involves applying traffic control measures in areas affected by an incident. Traffic management includes ensuring the availability of traffic control equipment and materials, knowledge of available fixed traffic control resources, and, most importantly, alternate route planning.

Clearance: Incident clearance is the process of removing the wreckage, debris, or any other element that disrupts the normal flow of traffic or forces lane closures and restoring the roadway capacity to its pre-incident condition.

Recovery: Recovery consists of restoring traffic flow to its normal condition at the site of the traffic incident. It encompasses the activities of site management, traffic management, and clearance. Traffic operations centers facilitate the recovery by managing the network-wide effects of traffic incidents and thus hastening recovery.

Figure 2-1 presents a timeline of the stages of the traffic incident management process. The traffic management and motorist information begin just after the detection of the incidents. To restore the normal flow after the incident, traffic management, including recovery methods, plays a crucial role.

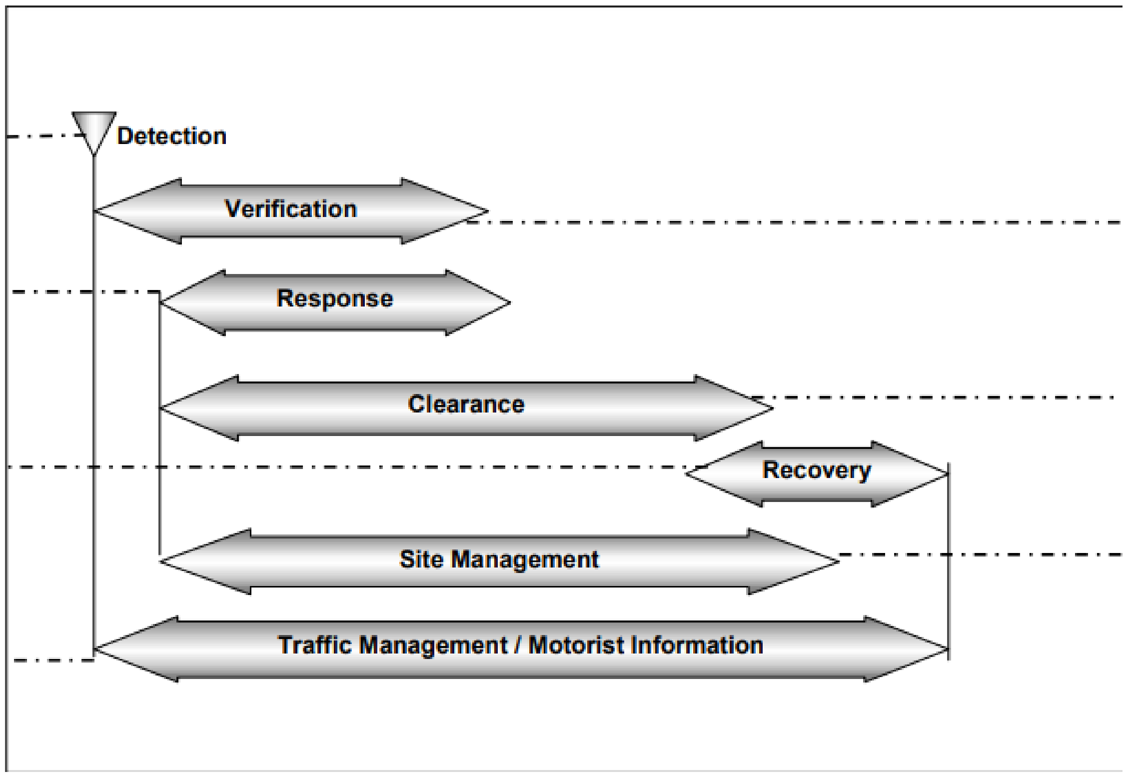


Figure 2-1: Timeline of Stages in the Traffic Incident Management Process (Source: Neudorff, et al., 2003)

A study by Martin et al. (2011) on freeway incident management using Vissim simulation in Utah identified the following factors for successful traffic management during the incident.

i. If the location of the incident does not allow a good choice of alternate routes, providing only information to the drivers does not help reduce delay; some other management measures can help in that case.

ii. Alternate routes along freeways during incidents offer lower travel times and lower network-wide delays compare to the incident impacted routes.

iii. If the alternate routes are arterial streets, traffic management strategies should be applied to these streets, especially for major freeway incidents (duration more than 60 minutes and level of closure greater than 50%).

A detour operation is a process that encourages traffic to take alternative routes to avoid incidents. Ng et al. (1995) analyzed a survey result which showed that the most common reason for choosing an alternative route is a crash (86% of the respondents to the survey), followed by road construction (79%), high traffic demands (71%), and time-saving by rerouting (49%). Understanding the importance of alternative routes, agencies have developed detour operation guidelines to better manage traffic considering diversion.

2.3 Detour Operation

Alternate route plans support the agency's goal of improved mobility and safety. An alternate route provides additional capacity to the primary route traffic and allows vehicles to circumvent the congested location. A study by Lin and Kou (2003) validated the importance of alternative route operations in response to a major freeway incident in terms of travel time benefits.

The Alternate Route Handbook (2006) provides comprehensive and general guidelines on planning and executing the detour operations involving various stakeholder agencies. According to this document, the key factors to consider for detour operation include the incident duration, the number of lanes blocked, the observed traffic conditions, the time of day, the day of the week, the capacity of the proposed alternative routes, and traffic status (Dunn Engineering Associates, 2006). Table 2-2 summarizes the criteria used in 2006 in several states to decide whether to execute the predeveloped alternate route plan.

Table 2-2: Criteria for Deciding Detour Operation (Source: Alternate Round Handbook, 2006)

Agency	Criteria
North Carolina DOT - Main Office	Complete closure of the highway in either direction is anticipated for fifteen minutes or longer.
North Carolina DOT - Charlotte Regional Office	No action or discussion occurs until fifteen minutes after the incident. An alternate route plan is deployed only after fifteen minutes if the highway is completely closed (all lanes closed, including the shoulder), and closure is expected to last at least an additional fifteen minutes (thirty minutes total).
New Jersey DOT	Level 1: Lane closures on a State highway that are expected to have a prolonged duration and impact on traffic. Level 2: Complete closure of a highway that is anticipated to last more than ninety minutes.
Oregon DOT	Incident with two or more lanes blocked, or incident with one lane blocked and expected to last more than twenty minutes.
New York State DOT Region-1	Detour operation will be implemented only when the highway is completely closed. It will not be implemented if at least one lane (or even the shoulder) is open.
Florida DOT District IV	Two or more lanes are blocked for at least two hours.
ARTIMIS (Ohio/Kentucky)	This plan has a detailed table with four different table levels, based on some different criteria, such as: • During the morning and afternoon peak hours, an advisory alternate route is deployed in the event of a two-lane closure for more than two hours or closure of more than two lanes for less than thirty minutes. • Mandatory alternate routes are deployed during peak hours when more than two lanes are closed for at least thirty minutes.
Ada County, Idaho	This plan sacrifices different levels of severity, including: • Levels C and D require the implementation of a diversion route. • Level C is an incident taking thirty to 120 minutes from detection to full recovery of the traffic flow. • Level D is an incident taking over two hours from its detection to full recovery (including full freeway closure in one or both directions).
Wisconsin DOT (Blue Route)	An incident that causes delays and exceeds thirty minutes.

Besides the Alternate Route Handbook guidelines, the Manual on Uniform Traffic Control Devices (MUTCD) (2009) states that major and intermediate incidents lasting more than thirty minutes usually require traffic diversion or detouring road users due to partial or full roadway closures. It also states that traffic diversion may not be necessary for minor incidents usually cleared within thirty minutes (FHWA, 2009). A study by Luo and Liu (2011) on detour decisions during freeway incidents indicated a timely and well-justified detour decisions could yield substantial benefits to both the driving populations and the entire community. The study identified the benefits of detouring operation even for minor incidents with relatively light volumes on both the freeway and detour routes.

Although the above guidelines provide the agencies basic information for setting detour operations, including selecting alternative routes, these guidelines are static. They do not consider the many dynamic variables that influence the decision to divert and select alternative routes in real-time operations. In general, it can be concluded that the incident attributes that significantly affect the traffic operations as well as induce the diversion of the traffics are:

i. number of lane blockage,

ii. duration of the incident

iii. time of day,

iv. day of the week,

v. location of the incident,

vi. traffic status on both affected and diverted routes,

vii. available capacity in the diverted route, and so on.

2.4 Advanced Traveler Information Systems (ATIS) to Support Diversion

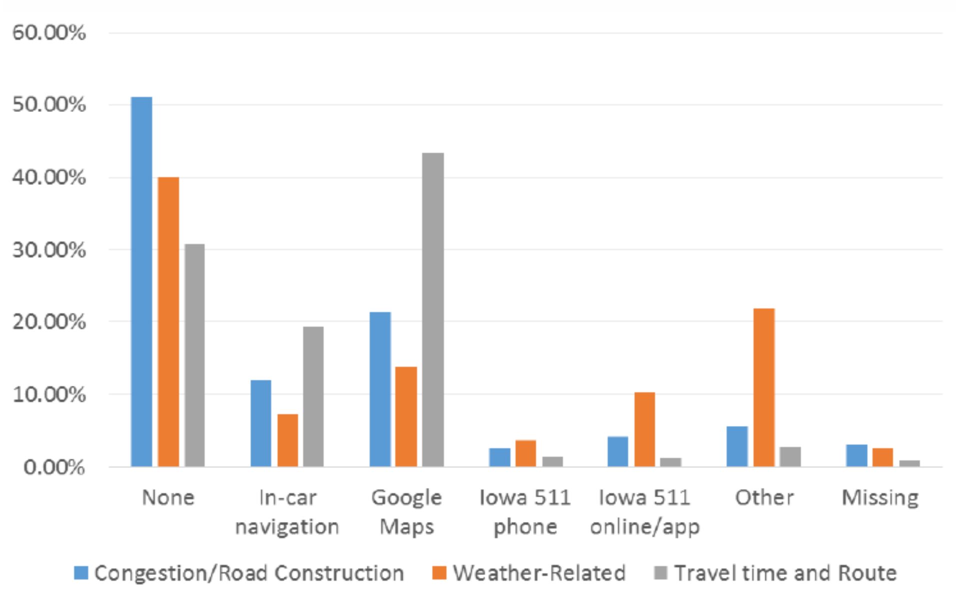
The diversion of the traffic during incidents is voluntary based on the traveler's utilities of the subject facility and the alternative routes. ATIS plays a significant role in the diversion mechanism by providing route information, and in some cases, diversion guidance to the users. Receiving information in real-time and the quality of the provided information will influence the diversion percentage.

With the advancement of ITS technologies and the associated deployment, most travelers are able to get real-time information about the current states of the system. Transportation agencies display incident and, in some cases, travel time information via dynamic message sign (DMS). Various studies showed positive impacts of DMS deployment during an incident in terms of reductions in link travel times (Ozbay et al., 2009; Yu et al., 2010). Agencies also broadcast information through Highway Advisory Radio (HAR). A study conducted by Sandt et al. (2017) for the Florida Department of Transportation (FDOT) assessed the impacts and usefulness of HAR in providing traveler information. The benefit-cost analysis estimated a benefit-cost ratio of 1.19, assuming that a 10% diversion was caused by HAR messages. The ratio was estimated to be 11.91 if the diversion is 100%.

Although the 511 phone system is deployed in most states, this system still lacks publicity and utilization. In a survey conducted in 2012, it was found that most of the respondents were unaware of the 511 system (Robinson et al., 2012).

Besides the agency's efforts, the travelers increasingly use private sector smartphone-based apps like Apple Maps, Google Maps, Waze Maps, etc., that provide them information about the traffic, routing, and also incident information before reaching the incident location. According to recent survey data, 85% of Americans own an internet-enabled handheld device, and most adults (as high as 93%), ages 18 to 64 years, commonly access the internet on mobile devices. These survey data indicate that mobile-internet access is becoming more important in people's daily lives than traditional media (e.g., TV, radio stations, and PC-based web applications).

A survey conducted for the Iowa Department of Transportation (Sharma et al., 2015) also identified the relative use of different traveler information services, as shown in Figure 2-2. The figure indicates a high number of Google Maps users.



Other: TV/News (2%, 11.18%, 0.59%); Radio (1.88%, 2.59%, 0.12%); Phone apps other than 511 (0.82%, 4.24%, 1.06%)

Figure 2-2: Traffic Information Services Used by the Survey Respondents (Source: Sharma et al., 2015)

2.5 Estimation of the Diversion of Traffic

A large number of studies have been performed to estimate the diversion of traffic. However, the question of how information affects traveler decision-making regarding routes, modes, and trip times under various operating conditions is still in search. Several researchers have used the stated preference approach to determine the percentage of travelers changing trip decisions in response to information disseminated by ATIS technologies. Based on this type of survey, studies concluded that the disseminated information could result in up to 40% to 70% of freeway traffic exiting the freeway ahead of an incident location (Madanat et al., 1995; Chatterjee et al., 2002; Al-Deek et al., 2009). It has been reported based on revealed preference survey results that the actual diversions are lower than those estimated based on stated preference surveys. However, information regarding the actual diversions due to traveler information remains limited. An extensive survey in six European countries to examine the impact of DMS on traffic diversion was conducted and found with the driver questionnaire results that the diversion rates are 0% to 7% for incident messages and 0% to 35% with route guidance information (Chatterjee and Mcdonald, 2004).

Using traffic detector data, researchers estimated the diversion to be in the range of 5% to 18% (Deeter, 2012; Haghani et al., 2013; Foo and Abdulhai, 2006). Hadi et al. (2013) conducted a study on the diversion of traffic in incident conditions based on detector data and found that the diversion rate ranges from about 8% for one out of five-lane blockages to about 25% when four out of the five lanes were blocked. Tariq et al. (2019) developed a method to predict the diversion due to incidents based on freeway mainline detector data combined with incident data. The developed method utilized a combination of cumulative volume analysis based on detector data, clustering analysis, and predictive data analysis. The results from applying the method suggest diversion from 1%-25% as a function of different incident attributes.

The above studies based on sensor data analyses revealed a 5% to 25% traffic diversion during incidents. The diversion of the traffic during the incident on the freeway triggers the need for the development of plans to accommodate the traffic on the alternative routes. The sudden increase in the traffic on the arterials causes congestion at traffic signals. The traffic signals typically run using plans based on time of day. The plans are developed using typical traffic patterns that do not consider the impact of additional traffic.

2.6 Coordination among the Routes during Diversion

Several efforts have been made to reduce the impact of the incident by coordinating the operations of the impacted route with the neighboring routes. Intuitively, disseminating traffic on the entire corridor, especially to the underutilized arterials, will reduce the overall impact of incidents. Researchers have been working to establish methodologies to realize the benefits of this tactic.

Under most incident scenarios, if proper diversion plans can be implemented in time, motorists can circumvent the congested segments and best use the available corridor capacity (Kim et al., 2017). After implementing the detour plan, specific coordination strategies are necessary to align with the detour option. The coordination operation combines both signal timing and diversion operation.

An earlier study used a hypothetical network to assess the effect of coordination on the impacted route and diverted route. Taylor and Narupiti (1996) used the NETSIM simulation program to examine the effectiveness of traffic diversion and signal timing modification for various incident conditions. A hypothetical surface street network was used in the simulation. The route impacted by the incident and the diverted route were coded in the simulation tool. Instead of fixed time control on the diverted route, the SCOOT adaptive signal control was simulated to facilitate the diversion. The traffic was forced to reroute at a selected intersection in the simulation. The adaptive signal timing on the diverted route reduced traffic delay compared to the fixed time signal timing. The study results demonstrated the effectiveness of the strategy increases with the increasing severity of the incident.

The University of Maryland developed a simulation-based learning environment called SimPLE (Simulation Processes in a Learning Environment), a dynamic simulation and visualization tool that allows users to see the effect of freeway lane closing and traffic diversion to an arterial road. An interactive signal timing system was built to adjust the green band and assess the arterial delays and total delays (Plaisant et al., 1998).

To accommodate the detour traffic on the arterials from the incident impacted freeway, Liu et al. (2011) proposed a model which assumes that diversion occurs using immediate upstream off-ramp from the incident and get back to the freeway in the next on-ramp beyond the incident, irrespective of the incident duration and severity and time of day.

Zhou (2008) used CORSIM microsimulation to evaluate the effects of incident management along with signal timing modifications for the parallel corridors. The analysis result implied that the percentage of the diverted traffic volume has a significant impact on the total delay of the entire network. A 10% diversion rate from the impacted freeway to the adjacent arterials was found to yield the minimum network-wide delay.

Traffic diverted to the alternative routes from the freeway through off-ramps, but some diversion may occur before reaching the impacted freeway facility. Managing the operations of the off-ramps reduce the chance of spill back to the freeway and encourage diversion to alternative routes. A study by Tian et al. (2002) indicated that traffic-responsive ramp metering and adaptive diamond interchange signal control produced effective system operations during freeway incident occurrences. A dynamic signal priority control strategy can mitigate the off-ramp queue spillback along with the reduction in delay, the average number of stops, and increased speed (Yang et al., 2014). The coordinated control strategies reduce travel time by about 8-25% for different scenarios during the incident compared to the existing condition (Jacob and Abdulhai, 2010).

The CFA Operation Handbook (2006) provides guidelines for efficiently managing traffic operations on the freeway and arterial streets. The empirical studies in four different cities (Glasgow, Seattle, Anaheim, and San Antonio) showed that CFA operation played a decisive role in reducing traffic congestion. For example, in Anaheim, CA, the implementation of alternative corridor operation plans (signal timing plans, ramp metering plans, DMS messages, and route diversion plans) during nonrecurring congestion reduces travel time by up to 30% (Urbanik et al., 2006).

The idea of CFA operation was further extended to a broader area under the umbrella term of Integrated Corridor Management (ICM), a promising tool in the congestion management toolbox by optimizing the use of existing infrastructure assets and leveraging unused capacity along the urban corridors. The ICM initiative harnesses the intelligent transportation system's potential to efficiently and proactively manage the congestion, hence improving the mobility of the goods and people in the corridor level.

The advancement of the ITS technologies, improvement of the TSMO program, associated development among the agencies, and availability of supporting data make ICM both practical and feasible. Eight pioneering sites were selected to establish the methodology and to assess the benefits of ICM. The pre-deployment assessment of ICM in three pioneer site corridors (i.e., U.S.-75, Dallas, Texas; I-15, San Diego, California, and I-394, Minneapolis, Minnesota) established the overall improvement of travel times, with improvements increasing nearly tenfold under conditions of high demand and severe traffic incident (Alexiadis and Armstrong, 2012). It was also observed that travel time benefits were concentrated in the vicinity of the incidents disrupting the flow in the peak direction, with travelers directly affected by the incident would experience the most significant benefits (Alexiadis and Chu, 2016).

2.7 Travel Time Prediction

As part of the study, a model is developed for predicting travel time in real-time using traffic, travel time, and incident data. The existing literature on travel time prediction using data analytics and machine learning is discussed in the following section.

2.7.1 Travel Time Prediction during Normal Condition

Chien and Kuchipudi (2003) developed a model using real-time and historical travel time for predicting travel time. A five-minute interval was chosen for the short-term travel time prediction. The Kalman filtering algorithm was applied for travel time prediction because of its significance in continuously updating the state variable as new observations become available. The reliability of predicted travel time depends on the uniformity of the traffic condition throughout the network. The model failed to predict the travel time in the cases of congestion and incidents.

Liu et al. (2006) utilized a State-Space Neural Networks (SSNN) model in combination with the Extended Kalman Filter (EKF) to predict the travel time. The developed model showed the ability to address special traffic patterns not included in the training data. Historical travel time data, along with traffic volume, was used to predict the travel time. Travel time was predicted for a five-minute horizon. The prediction model produced a root mean square error percentage (RMSEP) in the range from 16.5% to 21.2%. However, under extremely congested conditions (higher travel time values), the RMSEP values are higher because the experienced travel times are longer than the five-minute calculation time step.

He et al. (2010) proposed a Genetic Algorithm (GA) improved Back Propagation (BP) neural network for predicting the normal day's travel time in the freeway section. Normal travel time and traffic volume were used as the input parameters in the study. The model adopted a 3-layer BP network with one hidden layer. The monthly average relative error estimated by the model was 3.2%.

Shen (2008) predicted short-term freeway travel time using Dynamic Neural Networks (DNN) based on traffic detector data. Three different DNN topologies were considered in the study, and found that Time-Delayed Neural Network (TDNN) works better in the prediction. The prediction result was also compared with a simple multilayer perceptron (MLP) neural network and found a slightly better result. Speed and volume data were used in the study for predicting travel time at 15-minute intervals. The mean absolute percentage error (MAPE) when using the TDNN for the 15-minute prediction ranged from 22% to 26%.

Chen et al. (2013) proposed a method to predict the dynamic travel times using historical traffic pattern data. A 5-minute average speed data was used as the historical status of the traffic. The proposed method seeks historical candidates with similar traffic patterns to the current conditions. Afterward, the future traffic state was predicted by the subsequent traffic state of each candidate. A Euclidean distance measure was used to measure the dissimilarity between the historical and current status of the traffic. The measured time steps are 10-minute intervals till 60 minutes. The MAPE for the proposed algorithm was 5.96% to 6.37%, which was 15% better than the K-nearest neighbor (KNN) methods.

A Deep Belief Networks (DBN) model was proposed by Siripanpornchana et al. (2016) to predict travel time. In the proposed method, a stack of Restricted Boltzmann Machines (RBM) was used to learn generic traffic features in an unsupervised fashion automatically, and then a sigmoid regression was used to predict travel time in a supervised fashion. The MAPE of the prediction was within 7% of the actual values on most of the routes.

Mousa and Ishak (2018) developed travel times estimation model in the freeway network at five-minute intervals using Basic Safety Messages (BSM) data. They developed an eXtreme Gradient Boosting (XGB) model for short-term travel time prediction on freeways. The developed XGB tree-based ensemble prediction model was able to predict travel time accurately and consistently with variations during peak periods, with mean absolute percentage error in prediction about 5.9% and 7.8% for 5-minute and 30-minute horizons, respectively.

Zhao et al. (2018) proposed an optimal KNN for travel time prediction where data sparseness problem exists. Virtual sensor nodes were set up between the real Remote transportation microwave sensor (RTMS) nodes and the two-dimensional linear interpolation, and the piecewise method was applied to estimate the average travel time between two nodes. The state vector used in the prediction of travel time was Speed, volume, time of the day, and the traffic congestion level. Before completion and after completion of the missing data, the MAPE is 7.4% and 3.1%, respectively.

Duan et al. (2016) used long short-term memory (LSTM) neural network to predict the future step travel time using current state travel time. A 15-minute time interval was used to predict the multi-steps ahead travel time for 66 links in Highways England. Evaluation results showed that the 1-step ahead travel time prediction error is relatively small; the median of mean relative error for the 66 links in the experiments is 7.0% on the test set.

Liu et al. (2017) established a short-term travel time prediction model using historical 90 days travel time using LSTM combined with DNN. LSTM-DNN model was then tested along with linear regression, Ridge and LASSO regression, Auto-Regressive Integrated Moving Average (ARIMA), and DNN models under ten sets of sliding windows. The LSTM-DNN model showed the lowest MAPE for a sliding window of 60 minutes.

2.7.2 Travel Time Prediction during the Incident

Domenichini et al. (2012) estimated the travel time during the incident adding the travel time under normal traffic conditions, the time required to clear the incident, and time spent in the queue. The time required to clear the incident was estimated by analyzing historical incident clearance time using the Classification And Regression Tree (CART) model. Time spend in the queue was estimated using the Modello Clessidra (hourglass model). The model was able to predict the travel time during the incident to some extent with reasonable accuracy. However, the model required expert knowledge to consider the high variability of the incident.

Tatomir et al. (2009) estimated the travel time in the incident condition using the Ant-Based Control (ABC) Algorithm. The algorithm was used to estimate the travel speed in case of the incident, and based on the speed-density relationship, travel time was estimated. The typical day travel time was estimated using the historical travel time data. A 5-minute time interval was used to estimate the travel time dynamically. In the case of multiple segments with different speed limits, an average speed limit was assumed.

Yang and Qian (2019) used four different regression methods, i.e., Auto-Regressive Moving Average (ARMA), LASSO Linear Regression, Stepwise Regression, and Random Forest, to predict travel time 30-minute ahead-considering weather and incident. A correlation analysis was applied to select the important features that are highly correlated with the travel time of the target segments. K-means clustering analysis was performed to separate the data for sessional effect. The lowest average Normalized Root Mean Square Error (NRMSE) was 17% for Random forest, while the baseline ARMA model found an error of 38.4% for 30-minute ahead prediction.

Xia et al. (2010) predicted travel time in the freeway during the incident using a combination of detector data as well as incident data. During the incident, queuing analysis was performed to estimate the travel time, while the duration of the incident was estimated using a lookup table created from the historical incident duration data set. The predicted travel time showed reasonable accuracy. During an incident, travel time considering the incident's attributes produced better prediction than without consideration of the incident.

2.7.3 Summary of Travel Time Prediction Methods

In past research, statistical modeling, data analytics, and machine learning methods were applied to predict travel time on the highway facilities. Among these methods, with the increasing availability of data, the machine learning model has been widely used and showed better capability in capturing features of the time series data with higher accuracy. However, most of the existing studies do not consider traffic attributes of the subjected roadway or its surrounding roadways and incident attributes in predicting travel time. All these attributes have an important influence on the travel time of the roadway. So there is a need to include these attributes in the analysis and evaluate the performance of the method in the prediction of travel time estimation.

2.8 Modeling of Incidents and Diversion

Modeling incidents and resulting diversion in the alternative routes require considering various factors such as incident characteristics, traffic status on the affected facility and alternative routes, signal plan on the alternative routes, time-of-day, origin-destination, and so on. These time-varying and complex phenomena can be modeled utilizing a multi-resolution modeling (MRM) framework. MRM framework, which combines three different modeling levels, i.e., macroscopic, mesoscopic, and microscopic modeling, addresses issues beyond the capabilities of macroscopic models, mesoscopic models, and microscopic models by themselves (FHWA, 2012). The dynamic traffic assignment (DTA) technique in the mesoscopic level of this framework has the capability to simulate time-varying traffic diversion during incidents, whereas the microscopic level provides the necessary details of the traffic stream to develop signal control strategies. Since the incident induced diversion is a dynamic event and influenced by many factors such as incident attributes, traffic conditions, advanced traveler information systems, alternative route conditions, and signal status (Al-Deek et al., 2009; Khattak et al., 1992;

Tariq et al., 2019), an MRM framework can play a role in predicting diversion to support the traffic management during incidents.

In most implementations, MRM links a mesoscopic-based DTA model to both the regional travel demand models (macroscopic) and localized high-detailed models (microscopic). The MRM framework was successfully applied to support and ATM strategies, including during incident on arterials (Massahi et al., 2019), managed lane operations, integration of signal timing estimation modeling (Fakharian Qom, 2016), and DTA (Zlatkovic and Zhou, 2015), integrated active traffic operation evaluation (Mirchandani et al., 2018), and so on. This study utilizes MRM to identify the path-level diversion scenario demands and associated impacts on the alternative routes during incidents on the freeway.

2.9 Traffic Control during Incident

A smartly designed traffic signal control system is crucial to the efficient movement of the traffic in the arterials. Advancement of the system has the capability of reduction of traffic collisions, vehicle travel time, delay, parking, and energy consumption significantly. Coordination between the freeway and arterial during the incident requires signal timing in the arterial intersection in harmony with the sudden change of the traffic volume in the arterial. Apparently, signal timing should work in a fashion that arterial could accommodate additional diverted vehicles without deteriorating existing vehicular movement.

Transportation agencies usually operate the signal control systems based on time-of-day (TOD) plans. These plans are prepared using historical traffic flow data at different times of the day. The TOD schedule determines what time a plan will be active. The simplest schedules typically define an a.m., off-peak, and p.m. peak for weekdays and a different set of plans for weekends. The TOD plans cannot respond to the traffic variation that occurred due to incidents, which leads to high congestion and longer recovery times in the networks. With the increasing emphasis on active traffic management (ATM), some agencies have employed expert operators at traffic management centers (TMCs) to manage the traffic signal control actively during non-recurrent events (Tariq et al., 2020). The decisions made by such operators, however, are still reactive. Moreover, it is challenging for the operators to select the best plans given the many changing parameters in real-time operations during non-recurrent events (Kim et al., 2014).

Adaptive traffic control systems (ATCS) have been developed and implemented to react to the inherent traffic variations occurring from cycle to cycle, thus operate more efficiently than TOD-based systems. It has been reported that ATCS can reduce the delay during the incident conditions (Chilukuri et al., 2004). However, the ATCS performance under sudden demand surge and when the signal intersection approaches have long queues are uncertain. The oversaturation of an intersection or a movement negatively affects the performance of the ATCS system and may result in under allocation of green times to critical oversaturated movements (Campbell and Skabardonis, 2014; Lidbe et al., 2019). In addition, the existing ATCS systems deal with only the current traffic conditions as measured by the traffic sensors; thus, they are still reactive systems.

2.10 Traffic Signal Optimization

The essential design elements of traffic signal optimization are phase sequence, cycle length, green split, and offset. Optimizing the elements alone or multiple based on the proper objective function will lead to the proper timing for the subjected intersection.

The signal timing performance is not limited to the one intersection; instead, it influences the whole corridor.

2.10.1 Objectives Function in the Optimization

The performance of the optimization program largely depends on the objective function that will be maximized or minimized. In-depth domain knowledge and the nature of the problem are necessary for choosing the objective function. The demand pattern (i.e., symmetric, asymmetric, undersaturated, oversaturated) in the subjected intersection also plays a crucial role in choosing the objective function. For example, the dissipation of queues and removal of blockages in the oversaturated condition should be prioritized over the minimization of travel costs (Roess et al., 2013).

The objective functions vastly considered in the previous literature are the minimization of delay, minimization of travel time, maximizing the throughput-minus-queue, maximizing the number of completed trips, maximizing the weighted number of completed trips, and the number of stops. In a study of the selection of the objective function, Hajbabaie and Benekohal (2013) ranked the weighted number of completed trips as the most suitable candidate for oversaturated conditions followed by maximization of throughput-minus-queue. However, the authors mentioned that finding the completed trip is very complicated and used scarcely.

Delay minimization is the most widely used objective function for signal timing optimization. Sometimes it is used alone, but sometimes in combination with other objective functions. Webster (1958) and Miller (1963) established a traffic signal timing model and calculation method to get a minimum average delay of vehicles, building the foundation for modern fixed-time traffic signal control. Some other studies that used delay minimization are Zhang et al. (2010), Lo and Chow (2004).

Throughput plays a crucial role in the oversaturated conditions; hence, throughput-minus-queue maximization attracted much attention as an objective function of optimization. Studies such as those of Abu-Lebdeh and Benekohal (2000), Girianna and Benekohal (2002), and Putha et al. (2010) maximized throughput-minus-queue.

Christofa et al. (2016) proposed a new perspective of signal timing optimization. They proposed a real-time signal control system that optimizes signal settings based on the minimization of person delay on arterials. The system's underlying mixed-integer linear program minimizes person delay by explicitly accounting for the passenger occupancy of autos and transit vehicles.

2.10.2. Multi-Objective Optimization

Some studies aggregated multiple objectives functions and performed as a single objective optimization (Liu et al. 2011), while other studies performed multi-objective optimization and developed a Pareto front for selecting the optimal solution. Sun et al. (2003) demonstrated the efficiency of optimizing the average delay and the average number of stops in a multi-objective optimization for stochastic traffic arrival of traffic. Kesur (2010) investigated and suggested the use of multi-objective optimization when there are numerous optimization variables as the method improved the optimization efficiency over the single-objective optimization. Ezzat et al. (2014) found better results for multi-objective optimization over single objective optimization for oversaturated conditions. Other studies that used multi-objective optimizations include (Chen et al., 2011; Ceylan, 2013; Stevanovic et al., 2013),

2.10.3 Optimization Techniques

Heuristic optimization (i.e., Genetic Algorithm, Simulated Annealing, Ant Colony Optimization, etc.) is one of the popular techniques used in the traffic signal control system optimization for its fast and near-optimal solution. The model with feedback control, if appropriately applied, can provide solutions that are good approximations to the optimal solution. Genetic Algorithm (GA), one of the popular methods highly used in many previous studies. There are few other algorithms also used in some works. The literature related to optimization techniques is discussed in the following section.

2.10.3.1 Genetic Algorithm (GA)

Foy et al. (1992) implemented a genetic algorithm (GA) to generate optimal or near-optimal intersection traffic signal timing strategies that yield the smoothest traffic flow with the least average automobile delay. Hadi and Wallace (1993) later introduced the GA in the TRANSYT-7F (T7F), a signal timing optimization software, to optimize cycle length, phase sequences, and offsets. The association of GA with T7F proved very promising in the optimization of signal phasing and timing.

Park et al. (1999) proposed and evaluated three different GA-based optimization strategies: throughput maximization, average delay minimization, and average delay minimization with a penalty function. The evaluation recommends a delay minimization strategy because of its applicability to both undersaturated and oversaturated conditions. Hypothesis testing results on these three strategies indicated that the GA-based program with average delay minimization produced a superior signal-timing plan compared with those produced by other GA strategies and the T7F program in terms of queue time for oversaturated conditions.

Girianna and Benekohal (2004) applied GA to coordinate signals to maximize throughput. The optimization results demonstrated a control strategy that avoided queue spillback. They further extended this method to coordinate oversaturated signals and found it effective.

Sun et al. (2003) choose nondominated Sorting Genetic Algorithm II (NSGA-II) (Deb et al., 2002) for its better performance over other multi-objective evolutionary algorithms. They investigated the application of the NSGA-II in solving the multi-objective signal timing optimization problem considering deterministic and stochastic traffic patterns and found it efficient. They also proposed the Pareto-frontiers regression function to evaluate the trade-off among various traffic signal design objectives.

Memon and Bullen (1996) compared the gradient-based algorithm, quasi-Newton algorithm, with GA for signal timing optimization. Compared with gradient-based algorithms, GAs are more general and can optimize the phase sequence. Beyond their generality, GAs are considered more efficient in terms of implementation. They found that genetic algorithms are more efficient and effective than the quasi-Network method for this real-time application.

GA was found successful in solving signal timing-based Network Design Problem (NDP). The optimum timings obtained are compared with those obtained from a method that does not consider rerouting using GA. A simulation-assignment model provides the junction delays based on which travel costs are calculated. Besides delays at signalized junctions, the model also enables the consideration of delays at unsignalized junctions (Teklu et al., 2007). Li et al. (2013) used GA to optimize the multi-objective algorithm during the incident in the freeway concurrently diversion of the traffic in the arterial.

2.10.3.2 Simulated Annealing (SA) & Others

Hadi and Wallace (1994) recommended Simulated Annealing (SA), a meta-heuristic optimization algorithm, to implement in the T7F to optimize cycle length, phase sequences, and offsets simultaneously based on the progression. The authors found that SA is a computationally efficient algorithm compare to GA at that time. They also found that the algorithm has the potential for optimizing signal phasing and timing for arterial streets as well as multi-arterial networks.

Traffic signal control in a 4-leg intersection was simulated, and optimum signal timing by maximizing traffic flow was tried to obtain using both GA and SA. The goal is to study how traffic flows in the solutions produced by GA and SA when the problem size increases. This work concludes that SA seems to find better solutions than GA in small search spaces generally and that SA and GA are comparable in larger search spaces (Burvall and Olegard, 2015).

Li and Schonfeld (2015) applied the hybrid SA-GA algorithm proposed by Adler (Adler, 1993) to improve the performance of a GA for oversaturated signal timings optimization. The idea of this hybrid SA-GA algorithm is to use the selection rule of Metropolis to improve the efficiency of the GAs selection and crossover. The hybrid model was applied to optimize an oversaturated arterial intersections signal timing and compared it with SA and GA results applied separately. However, the hybrid models required more central processing unit (CPU) time to reach an optimized solution because the solution domain is larger, and the evaluation time of each chromosome is longer.

Putha et al. (2012) employed Ant Colony Algorithm (ACA) algorithm in an oversaturated condition to minimize queues and remove the blockage in an alternative to

GA. They found that the GA algorithm is too slow or requires very high computational power to be useful in real-time situations. The result from two scenarios was compared for both ACA and GA and found that ACA is consistently more effective for a larger number of trials and to provide more reliable solutions.

In another study, He and Hou (2012) used ACA to effectively separate the conflict of the traffic flow and optimize the cycle time and saturation of an intersection. The performance indices: time delay, number of stops, and traffic capacity obtained from ACA were compared with Webster algorithm, GA. Numerical results showed that ACA is a feasible and straightforward method for signal timing optimization problems.

Chen and Xu (2006) used the particle swarm optimization (PSO) algorithm to solve the traffic signal timing optimization problem. A local fuzzy-logic controller (FLC) installed at each junction is used to provide some initial solutions for the particle swarm optimization algorithm. Membership functions and the rules of the fuzzy logic controller (FLC) were optimized using the particle swarm optimization (PSO) algorithm. The cycle length, the phasing splitting, and offset of all signals in the adjacent junction's system are optimized using the new PSO algorithm. PSO has been proved to be a competitor to the standard genetic algorithm (GA). Comparisons between PSO and GA were made with regards to the performance by Lampinen (2002), which points out that the PSO performs well in the early iterations but has problems in reaching a near-optimal solution in several benchmark functions.

Zhao et al. (2018) considered a non-dominated sorting artificial bee colony (ABC) algorithm to solve the multi-objective optimization model considering the reasonable balance between vehicle delay and stops of a fixed-time signal control parameter of unsaturated intersections. Experimental results showed that the convergence is better in the non-dominated sorting ABC algorithm than in the non-dominated sorting genetic algorithm II. Moreover, the algorithm can solve the Pareto front of a multi-objective problem, thereby improving the vehicle delay and stops simultaneously.

2.10.3.3 Optimization in Simulation Environment

Among the heuristic algorithms, only the GA was successfully applied in the traffic signal control system for commercial productions (Park et al., 1999). Conventional optimization methods such as integer programming, hill climbing, or descent gradient searching have been gradually overshadowed by genetic algorithms in many areas, including traffic signal operation (Lee et al., 2005). Initially, GA was used along with macroscopic simulation to optimize the signal. However, a recent study also revealed that a signal timing plan based on a direct signal optimization using a stochastic and microscopic simulation model produces better performance than that of a macroscopic simulation-based method (Rouphail et al., 2000). Park et al. (2001) applied GA along with CORSIM microscopic simulation to optimize the signal timing directly. The result indicated that a well-calibrated simulation program is crucial in the evaluation of signal timing plans.

Genetic algorithm optimizations of traffic signal timings have been shown to be effective, continually outperforming traditional optimization tools such as Synchro and T7F (Stevanovic et al., 2007).

Stevanovic et al. (2007) utilized GA optimization in the Vissim microsimulation interface to directly optimize the signal timing. The program called VISGAOST optimized four basic signal timing parameters with Vissim microsimulation as an evaluation environment. The program added new optimization features such as the optimization of phasing sequences, multiple coordinated systems, and uncoordinated intersections, fully actuated isolated intersections and multiple periods. The program was tested on a hypothetical grid network and a real-world arterial of actuated–coordinated intersections in Park City, Utah. Evaluating the result, it was found that timing plans optimized by the genetic algorithm outperformed the best Synchro plans in both cases, reducing delay and stops by at least 5%.

Branke et al. (2007) used Non-dominated Sorting Genetic Algorithm II (NSGA-II), a multi-objective optimization, for developing signal timing plans using VISSIM simulation (Branke et al., 2007). Vissim based NSGA-II algorithm was found successful in maximizing throughput and minimizing queue ratio in an oversaturated condition and produced a better plan managing the traffic flow efficiently than Synchro (Li et al., 2013).

Stevanovic et al. (2013) used Vissim based NSGA-II multi-objective optimization for signal plan development, considering mobility and safety together (Stevanovic et al., 2013). The authors used throughput as the measure of mobility, while the number of conflicts was used to measure safety. The result demonstrated a 7% decrease in conflicts while maintaining the same throughput compared to the initial level.

2.11 Connected Vehicle (CV) and High-Resolution Controller (HRC) Data

The application of machine learning for predictive modeling is very promising due to its ability to capture the non-linear relationship within the dataset. Besides the conventional data collection techniques such as loop detector, probe detector, microwave detector, and so on, there are two other data sources, i.e., CV and HRC data, gaining popularity for their high fidelity. In addition to its many other benefits, CV can act as a mobile sensor for collecting various traffic data in high resolution. Similarly, HRC data allows the estimation of various performance measures of the traffic signal. Several uses of the CV data and HRC data are discussed in Sections 2.12.1 and 2.12.2

2.11.1 Connected Vehicles (CVs)

Implementation of connected vehicle (CV) will accrue numerous benefits from both users' and agencies' end (i.e., capacity increase, delay reduction, crash reduction, and mobile data sources). Guler et al. (2014) observed that CV technology could significantly improve the operation of traffic at signalized intersections. Increases in the penetration rate from 0% up to 60% can significantly reduce the average delay. Different performance measures, i.e., travel time estimation, incident detection, signalized left-turn assist, and traffic volume estimation, can also be evaluated using the CV, which can replace the traditional way performance measurements (Iqbal et al., 2018; Mwakalonge et al., 2019). CV can work as a mobile sensor by continuously reporting their status to roadside equipment (RSE) through vehicle-to-infrastructure (V2I) communication. Therefore, CVs hold great potential to reduce or even eliminate the need for fixed-location detectors in the existing signal systems. When penetration rates are low, the CV data could be used to generate performance measures for fine-tuning traffic signals periodically based on the offline analysis. When penetration rates are high, it becomes feasible to operate adaptive signal control that solely depends on the CV input.

The conventional vehicle detector for signal operation can be superseded by CVs technology (Zheng and Liu, 2017). CV data can be used to estimate the traffic volume under low market penetration. Zheng and Liu (2017) developed an approach to estimate traffic volume using GPS trajectory data from CV or navigation devices. The arrival of vehicles at the signalized intersection was modeled as a Poisson distribution. The estimation problem is formulated as a maximum likelihood problem given multiple observed trajectories from CVs approaching the intersection. An expectation maximization (EM) procedure is derived to solve the estimation problem. The model was validated using the field data and found that the proposed approach can work as low as 10% market penetration.

Bekiaris-Liberis et al. (2016) developed a traffic state estimation method by estimating the percentage of CVs with respect to the total number of vehicles. In the model development, the density and flow of the connected vehicle were obtained from their regularly reported positions. A Kalman filter method was employed in the estimation schemes.

The real-time queue estimation is very crucial for a queue-based adaptive traffic control system as the traffic control allocates the green time using the queue. Tiaprasert et al. (2015) proposed a mathematical model for real-time queue estimation using CV technology. The model worked without signal timing, traffic volume, or queue characteristics as basic inputs. Simultaneously, the model worked with both fixed-time signals and actuated signals. The model was validated for both pre-timed and actuated controls with a Vissim microscopic simulator that showed that it could reasonably accurately estimate the queue for both control systems.

Christofa et al. (2013) developed two queue spillback detection methods based on CV or probe data. One method uses only CV data and is based on the assumption that non-equipped vehicles in the queue that arrive after the last CV-equipped vehicle can be modeled by using a geometric distribution. The second spillback detection method combines CV data with information about the upstream signal settings and is based on the kinematic wave theory of traffic. Both methods provide a good detection of queue spillback with reasonable accuracy.

Goodall et al. (2016) proposed an algorithm to determine the locations of individual conventional vehicles based on the behaviors of nearby communicating vehicles by comparing a communicating vehicle's acceleration with its expected acceleration as predicted by a car-following model. The algorithm can predict the locations of 30% of vehicles with 9-m accuracy in the same lane, with only 10% of vehicles communicating. Similar improvements were found at other initial penetration rates of less than 80%. The algorithm was estimated using the Vissim simulation data and validated using the NGSIM data.

Cao et al. (2019) developed a method to estimate left-turn queue spillback probability using CV data. The method initially estimated the total queue length using the last CV location that was finally used to estimate the left turn spillback. The developed method was validated using Vissim microscopic simulation. Based on the probability of forming queue spillback, a new signal timing based on the CV data was proposed. The results of the methods demonstrated a reduction in average delay by 20%.

2.11.2 High-Resolution Controller Data

High-resolution controller (HRC) data obtained from the controllers is event-based data with a temporal fidelity of 0.1 s. (Sturdevant et al., 2012). The data is recorded through a data logger software interface in the controller and captures all detection and phase events at a given intersection. The cycle-by-cycle data with such resolution provides numerous advantages, especially for the automatic performance measure of the signal.

In a collaborative effort among Econolite Control Products, Inc., the Indiana Department of Transportation (INDOT), and Purdue University, the Econolite ASC/3 controller software was enhanced to include a data logger to collect time-stamped phase and detector state changes (Smaglik et al., 2007). The obtained data has a specific format consisting of three columns: “Timestamp,” “Event Type,” and “Parameter.” These data can be used for various performance measures. The detailed formatting of the HRC data is discussed in Chapter IV. Several existing applications of the HRC data are discussed in the following section.

Smaglik et al. (2007) developed an integrated general-purpose data collection module within a National Electrical Manufacturers Association (NEMA) actuated traffic signal controller for the collection of cycle-by-cycle data. They used those data for producing quantitative graphs to assess arterial progression, phase capacity utilization, movement delay, and served volumes.

Liu and Ma (2008) successfully built a system called SMART-SIGNAL (Systematic Monitoring of Arterial Road Traffic and Signals). The system can simultaneously collect and archive event-based traffic signal data at multiple intersections and automatically generates real-time performance measures, including queue length, travel time, and the number of stops. The system collects two types of data, the signal event and detector event, and is stored in a log file every day. The system was installed on eleven intersections along France Avenue in Hennepin County, Minnesota. Based on the collected data, the authors proposed a mathematical time-dependent queue length estimation model under both under-saturated and over-saturated conditions. The model provided a reasonable estimation of intersection queue lengths and arterial travel times. The model was further used to evaluate a signal retiming effort on France Avenue.

Day and Bullock (2011) used HRC data to investigate the performance of various algorithms in the offset optimization problem. The HRC raw data were aggregated to obtain cyclic flow profiles, leading to estimates of delay and the number of vehicles arriving on the green. Two alternative objective functions, i.e., minimizing delay and maximizing the number of vehicles arriving on green (*Ng*) were used in the optimization. For both objectives, the outcomes of all algorithms were similar regarding their performance relative to the original offsets. The combination method, hill climbing, and genetic algorithms operated with similar computational efficiency.

HRC data provided a superior method to coordinate the traffic signals at the adjacent intersections based on the arrival on green. The developed method would be able to replace the existing decision to coordinate adjacent signals based on rules of thumb (e.g., one mile is considered a threshold upper bound link distance), or analysis of link volumes compared to distances or using traffic flow models (Day et al., 2011).

Day et al. (2011) developed a methodology for analyzing incoming traffic flow using high-resolution signal event data. Primarily, the upstream beginning of green time is projected forward in time to a downstream detector, which would be the upstream detector of the next intersection. By subtracting the upstream beginning of green time (plus a baseline travel time) from the detector arrival times, they developed a platoon profile by aggregating the data across successive cycles.

The quality of progression, an important performance measure of signal, can be done using HRC data by building Purdue Coordination Diagram (PCD) (Day et al., 2010;

Day et al., 2012). Visualizing the data in PCD helps identify the quality of progression; it also helps the agency infer the reasons behind poor progression (Day et al., 2011).

Liu and Hu (2013) developed a method using high-resolution controller data to assist ICM operation. They used the high-resolution controller data to optimize the signal control by maximizing the flow during the diversion of the traffic from the freeway. A microscopic simulation model was used to evaluate the performance of the method. The results demonstrated the method's effectiveness for reducing network congestion and improves network performance in terms of average delay per vehicle, the average number of stops, and average speed.

Dakic et al. (2017) developed a new high-resolution performance measure-Average Arrivals on Green Ratio (AAOGR) that consider the variability of cycle length and green time on a cycle-by-cycle basis and provides information on the ratio of vehicles that passed through the intersection per second of green time. Moreover, a discharge rate-based model was also proposed for computing approach delay, which considers the possible queue build-up during red and provides a better estimation of actual approach delay per vehicle if measured in front of the advanced detectors.

Recently, Tariq et al. (2021) developed a multi-objective optimization-based calibration methodology for the microscopic simulation model calibration using HRC data. The calibrated model produced significantly lower errors in the split utilization ratio, green utilization ratio, arrival on green, and travel time compared to a simulation model that uses the simulation model's default parameters.

In summary, the above studies revealed the potential of CV data and HRC data for various performance measures and the development of new methods and models to manage traffic better. This study explores new ways of use of CV and HRC data and developed methods based on the findings. The study develops a method for traffic congestion pattern identification.

2.12 Congestion Patterns Identification

Many studies in the literature identified the roadway congestion patterns based on macroscopic traffic attributes such as flow, travel time, speed, and so on. Zhang et al. (2016) developed a method to identify the traffic flow pattern as well as anomaly detection in the flow. A dictionary-based compression theory was employed to find the congestion pattern at the detector, intersection, and sub-region levels. Zhu et al. (2016) analyzed traffic flow data using Fuzzy C-Means clustering to identify the patterns. Based on the findings, it divided the intersection flow as high peak, evening peak, and flat peak.

Recently, vehicle trajectory data have become more accessible for identifying the traffic congestions in the network at a more detailed level, including both arterials and freeways. Jianming et al. (2012) developed a method that uses vehicle trajectory data as an image that used self-correlation to extract the propagation speed of the congestion wave. This extraction was used to construct a congestion template to identify the current congestion as well as intensity using a matching algorithm. Several other studies used taxi trajectory data to evaluate the congestion pattern utilizing machine learning techniques (Liu et al., 2015; Thianniwet et al., 2009).

Krishnakumari et al. (2017) proposed an image processing method using speed data contour to identify the congestion patterns. Using the expert knowledge, all the patterns found in the data set were manually classified into five classes prior to the training of classification models for use in real-time operations. The model produced an accuracy of 70% in the identification of the congestion patterns.

The above studies measured the congestion at the link level and cannot differentiate between congestions caused by different turning movements (e.g., left turn vs. through movement). For designing the proper phase splits at the signalized intersection, it is essential to ascertain the movement level congestion. Besides, the data used in those studies were aggregated data. This kind of coarse data may not be able to find the inherent reason for the congestions.

2.13 Summary

ICM has the potential to provide significant benefits in managing the congestion at the corridor level. It evaluates the entire corridor as a system rather than individual components in a coordinated fashion. Coordinated freeway and arterial (CFA) operation during the incident in the freeway is an integral part of the ICM. Existing studies related to the coordination strategies for both CFA and ICM were reviewed. The reviews of the existing TIM systems indicate that the management of alternative routes during the incident is based on fixed alternative routes for the diversion irrespective of the incidents and traffic conditions.

Signal control plans on the alternative routes as part of diversion route management were also reviewed. The reviewed studies included traffic signal control during incidents, objective functions selection, and optimization methods. In the plan development, microscopic simulation-based optimization was found effective in the literature. Studies also revealed the superiority of multi-objective optimization over single-objective optimization. Researchers demonstrated that the heuristics technique to solve the simulation-based optimization problem because of its quick and near-optimal solution.

This chapter also reviewed the use of data from two imminent technology, CV and HRC, in traffic management. In this regard, existing studies related to the use of CV and HRC data were reviewed. The most common use of CV data was found to estimate the traffic state and HRC data to measure signal control performance.

CHAPTER III

METHODOLOGY DEVELOPMENT

This chapter describes the methodologies utilized in the study to support the goal and objectives of this research. The first section provides an overall framework of the methods, and the subsequent sections illustrate each step of the methods.

3.1 Methodological Framework

The overall goal of the research is to develop models and associated methods for the management of diversion routes during freeway incidents using existing data sources and exploring the use of emerging data sources for this purpose. The methodology used in this research to achieve the goals primarily consists of three parts:

- Develop a method for predicting the critical routes utilized by motorists during incidents and the associated traffic and incident conditions under which the diversion occurs due to the incident on the freeway.
- Develop a method for generating special signal timing plans and estimate the benefits of activating those plans to mitigate the deterioration in the performance of the movements of critical route intersections due to diversion.
- Estimate the benefits of utilizing new data sources (i.e., CV, HRC) as part of the methodology to support traffic management in the network.

Figure 3.1 shows the basic architecture of the methodology followed in the study to achieve the overall goal. Subsequent sections discuss detailed steps of each part of the architecture.

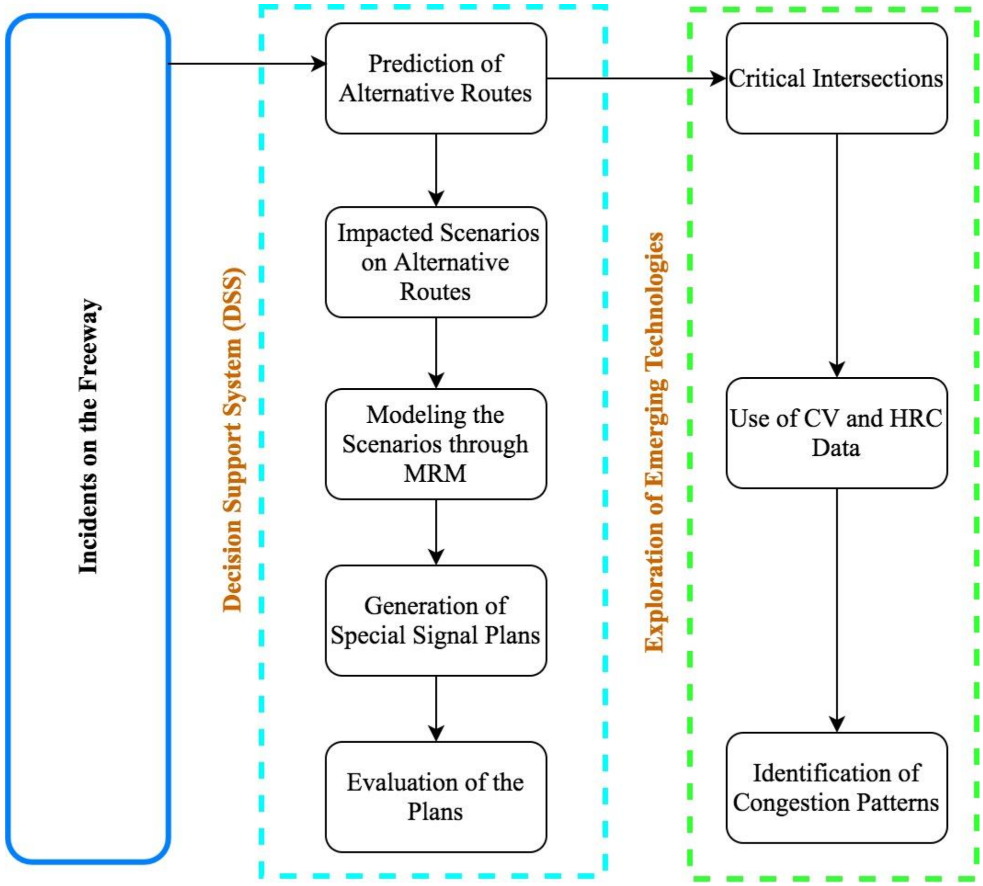


Figure 3-1: Methodological Framework

3.2 Prediction of Critical Diversion Routes

One of the crucial objectives of the study is the prediction of the critical alternative route(s) which motorists will use for diversion in the prediction horizon. Diversion is a dynamic phenomenon and requires real-time traffic and incident data in the network. For this purpose, a measurement termed Δ-Travel Time (ΔTT) is used in the study. It is defined as the percentage change in travel time on the potential alternative route in the prediction horizon after the occurrence of each incident with respect to the travel time of the base (normal) condition. Equation 3-1 shows the way of estimation of ΔTT.

$$\Delta TT = \frac{TT_i - TT_n}{TT_n} \times 100 \tag{3-1}$$

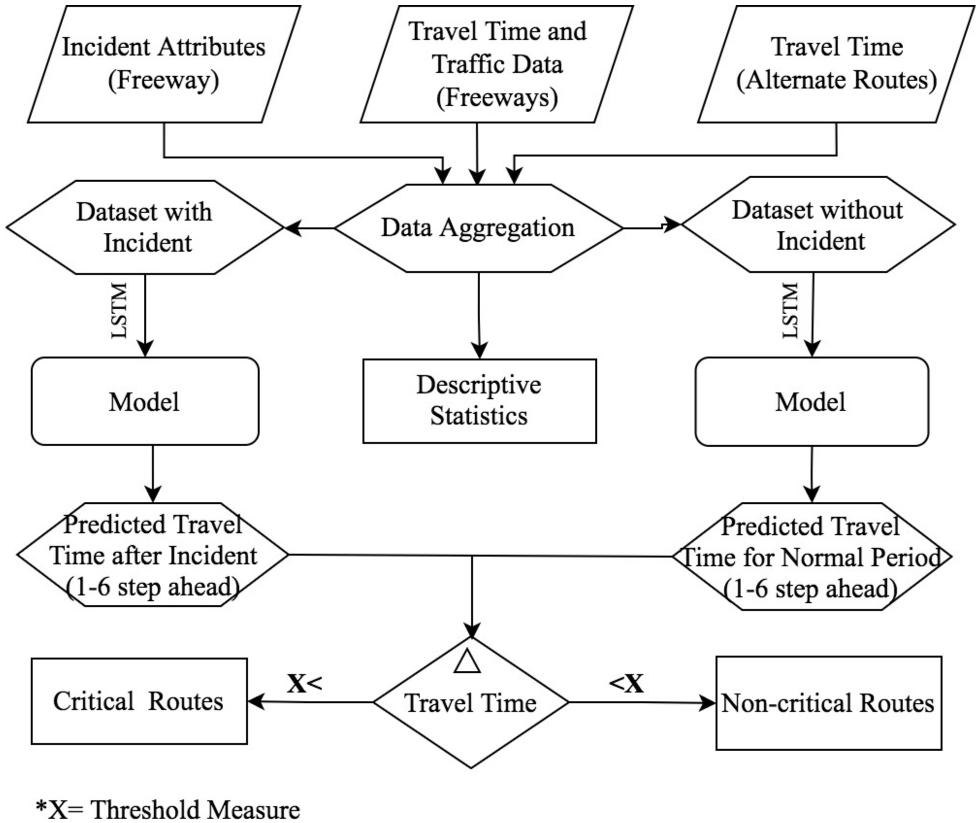
where

ΔTT = percentage change in travel time on the potential alternative route in the prediction horizon after the occurrence of each incident with respect to the travel time of the base (normal) condition,

TT_n= predicted travel time during the normal period, and

TT_i= predicted travel time during the incident.

The *ΔTT* acts as a threshold measure to identify the critical routes and the mobility impacts on these routes. Route(s) with higher *ΔTT* is considered as the critical route(s) as more vehicles will use those routes to divert from the freeway. In real-world applications, the threshold value of *ΔTT* can be set by the local agencies based on the local conditions. The methodology followed to predict the alternative routes in this study is shown in Figure 3-2.



LSTM=Long Short-Term Memory

Figure 3-2: Methodology for Predicting Alternative Routes

The estimation of *ΔTT* requires the estimation of travel time on the alternative routes in the analysis horizon for both incident and normal conditions. A model that can capture the dynamic features of the time series data is necessary to predict the travel time. The following subsections describe the method and associated models used in the study to predict travel time.

3.2.1 Methods for Travel Time Prediction

The literature review on travel time prediction methods revealed that machine learning methods could capture attributes of the time series data. Long Short-Term Memory (LSTM), a variety of RNN, has been successfully applied in many real-world problems involving sequence data such as music generation and speech recognition

(Hochreiter and Schmidhuber, 1997; Graves, 2012). The LSTM method was found very effective in capturing the time-dependent pattern of the traffic data. Prediction of the travel time of traffic using LSTM was shown to produce promising results (Duan et al., 2016). The method has been successfully applied to traffic performance prediction. It has achieved better performance in recent years (Ma et al. 2015) because of its better mechanics to model the traffic dynamics in road networks as it can model long-term dependence in time series and extract features from traffic data with recurrent feedback. LSTM is much faster to train than standard RNNs and MLPs and found slightly more accurate in frame-wise phoneme classification (mapping a sequence of speech frames to a sequence of phoneme labels associated with those frames) (Graves and Schmidhuber, 2005). Since LSTM is built on top of the basic principle of RNN, both methods are described here.

3.2.2 Recurrent Neural Network

The architecture of the RNN considers the time sequence relationship among the data. The basic structure of the RNN is shown in Figure 3-3 (Giles et al., 1994). The hidden layer between the input layer and output layer of each neural network is connected with the adjacent hidden layer of the following neural network in time sequence. Information retained in the former network hidden layer passes through the established connection to the later network. In Figure 3-3, the weight matrices, *U* and *W* between the layers, and the weight matrix *V* between neural networks are updated at each hidden state, h_t. However, the RNN itself fails to analyze the time series, as the back-propagating error vanishes or explodes exponentially with time (Hochreiter and Schmidhuber, 1997; Hochreiter et al., 2001).

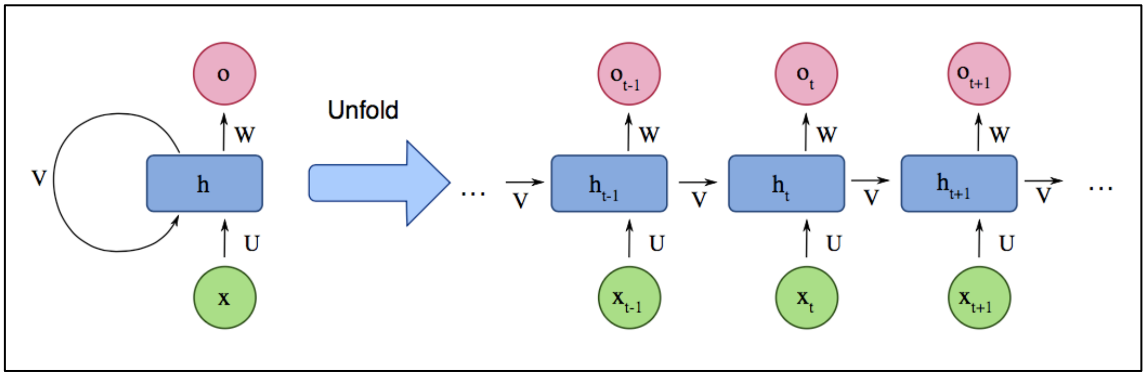


Figure 3-3: Basic Recurrent Neural Network (Source: Deloche, 2019)

3.2.3 Long Short-Term Memory

Long Short-term Memory (LSTM) is a recurrent network architecture along with an appropriate gradient-based learning algorithm (Hochreiter and Schmidhuber, 1997). The introduction of the gradient-based algorithm in the LSTM eliminates the exploding or vanishing gradient by enforcing constant error flow through the internal hidden layer called LSTM. The architecture of LSTM is shown in Figure 3-4, where an LSTM cell has three gates: forget gate layer, F_t input gate layer, I_t output gate layer, O_t. The forget gate layer decides which information from the cell state C_{t-1} at h_{t-1} will forget and which information will keep for feeding. The input gate layer will decide which new information will preserve that produce a candidate cell state $\hat{C}_t$. A new cell state C_t is produced at h_t in the output gate layer, combining both forget gate layer and input gate layer information.

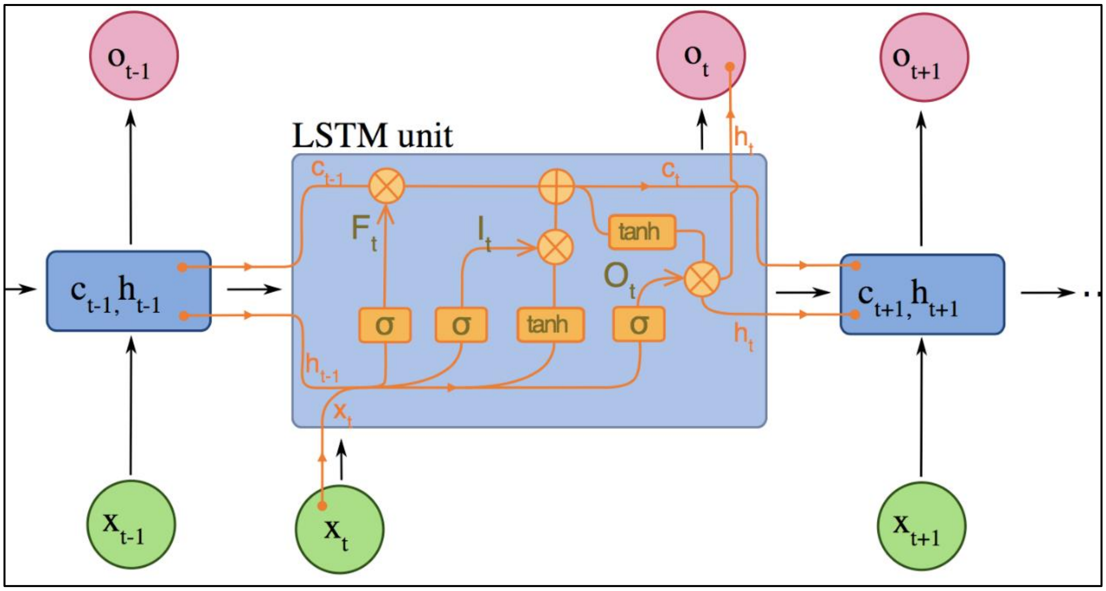


Figure 3-4: Long Short-Term Memory Unit (Source: Deloche, 2019)

The forget gate scales the internal state of the cell before adding it as input to the cell through the self-recurrent connection of the cell, therefore adaptively forgetting or resetting the cell's memory. Besides, the modern LSTM architecture contains peephole connections from its internal cells to the gates in the same cell to learn the precise timing of the outputs (Gers et al., 2003). An LSTM network computes a mapping from an input sequence $x = (x_1, ..., x_t)$ to an output sequence $y = (y_1, ..., y_t)$ by calculating the network unit activations using the following equations iteratively from $t = 1$ to T:

$$I_t = \sigma(W_{ix}x_t + W_{im}m_{t-1} + W_{ic}c_{t-1} + b_i) \tag{3-2}$$

$$F_t = \sigma(W_{fx}x_t + W_{fm}m_{t-1} + W_{fc}c_{t-1} + b_f) \tag{3-3}$$

$$c_t = F_t \odot c_{t-1} + I_t \odot g(W_{cx}x_t + W_{cm}m_{t-1} + b_c) \tag{3-4}$$

$$O_t = \sigma(W_{ox}x_t + W_{om}m_{t-1} + W_{oc}c_t + b_o) \tag{3-5}$$

$$m_t = O_t \odot h(c_t) \tag{3-6}$$

$$y_t = \phi(W_{ym}m_t + b_y) \tag{3-7}$$

where

W_{ic}, W_{fc}, W_{oc} = diagonal weight matrices for peephole connections,

b = terms denote bias vectors (b_i is the input gate bias vector),

σ = is the logistic sigmoid function,

I, F, O = input gate, forget gate, and output gate, respectively,

c = cell activation vectors,

m = cell output activation vector,

$\odot$ = element-wise product of the vectors,

g, h = cell input and cell output activation functions, and

$tanh$ = network output activation function.

3.2.4 Selection of Prediction Horizon

The prediction horizon is the time window after the incident occurrence, at which the prediction of the incident impact is made starts. It is the time used to account for the fact that the vehicles do not start diverting until a significant queue and delay are built on the freeway and that the vehicles take some time to reach the alternative route location to cause delay at the critical intersections. After the incident occurs on the freeway, the traveler needs some time to receive information about the increasing delay, react to the information, and reach the critical intersections.

From the signal timing plan point of view, it is not practical to change the plan very frequently because the transition between plans takes some time, and frequent switching between the plans, if not implemented correctly, can cause deterioration in traffic performance (Alexiadis and Chu, 2016). Besides, the proposed system can be classified as a "proactive" traffic responsive system, in which the utilization of a 15-minute interval is common. Adaptive signal control works with higher time granularity, but such systems have not been proven to work with surges in demands or abrupt reductions in capacities. Considering all these issues, a 15-minute prediction horizon from 15-90 minutes of the incident is selected for use in this study,

3.2.5 Model Training and Validation

Training and validation of the model, including the configuration of the hyperparameters of the model, is a crucial step in model development. The Python programming language with Keras library and Tensorflow library in the backend are used to train and validate the LSTM model. The model is derived to predict the travel time in the prediction time horizon. The input variables used in the model are the timesteps, time-of-day, incident severity, number of lane blockage, the location of the incident (distance from a reference point), average speed and average volume per lane on the freeway, and the travel times of both the freeway and arterials. The model provides predicted travel times for all alternative arterial segments for the prediction time horizon, divided into six steps (Steps 1 to 6 represent the 15-minute interval from 15-90 minutes after the incident). It is trained for predicting the travel times for the intervals between 15-90 minutes after the occurrence of the incident, where every 15-minute data points are used to predict the next 15-minute travel times.

The models' hyperparameters are fine-tuned, considering the resulting loss function targeting the achievement of better results in the prediction. Three hidden layers are used from the usually utilized range of two to five hidden layers besides the input and output layers. Inputs are fed in five batches for 100 epochs. The adaptive moment estimation (Adam) optimizer is used with a learning rate of 0.0001. Since the input variables consist of both categorical and continuous variables, one-hot encoding is used to convert the categorical variables (i.e., time-period after the incident, time-of-day, incident severity, number of lane blockage) into dummy variables for use in the model. The dataset is randomly divided into two parts: a training set and a test set, with the training set consisting of eighty percent of the data points and the test set consisting of twenty percent of data points. The model is trained using the training data set; its accuracy is estimated using the test data set.

3.2.6 Accuracy of the Travel Time Prediction

The accuracy of the model for travel time prediction during incidents utilizing the nonlinear relationship among traffic parameter inputs, incident attribute inputs, and travel time is determined using error functions of the test data sets. Two different performance measures: the Mean Absolute Error (MAE) and the Mean Absolute Percentage Error (MAPE) are estimated using Equations 3-8 and 3-9, respectively.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - y| \tag{3-8}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - y}{y_i}\right| \tag{3-9}$$

where y_i is the predicted travel time for i^{th} observation, y is the ground truth travel time, n is the total number of observations.

The MAE performance measure examines the cumulative deviations of the predicted outputs from the targets, while the MAPE translates the deviations into percentage forms by comparing them with the absolute target values. The MAPE is a measure of accuracy.

3.3 Special Signal Plan Generation

After identifying the critical alternative routes utilized by motorists for diversion, the next step is to develop and implement the proper signal timing in those critical alternative routes to facilitate the diversion. The process includes two major steps: i) estimation of path-level traffic demand and ii) optimization of the signal plan.

3.3.1 Estimation of Path-level Traffic Demand

The estimation of path-level traffic is necessary for each diversion scenario as it is one of the crucial inputs for designing the signal timing along the diversion path. Wrong estimation of the traffic demand on the intersection deteriorates the performance of the intersection and the corridor as well. In this study, the path-level traffic demand is estimated using the following two steps in combination: a) identification of the impacted scenarios and b) multi-resolution modeling.

3.3.1.1 Identification of Impacted Scenarios

The sudden demand surge of traffic on the alternative routes due to diversion deteriorates the performance of the routes. The higher is the traffic diversion, the higher is the expected impacts on the travel time of the corresponding routes. These travel time changes, previously termed as ΔTT, are further used to identify the extent of the effect of diversion under different representative scenarios. The representative scenarios are identified using clustering analysis on the ΔTT of the potential alternative routes for all six 15-minute intervals. K-means, a widely used clustering method, is utilized to identify the scenarios. Incidents responsible for those scenarios are identified from the data set to measure the path-level demand. The algorithm and features of the K-means method are described in Section 3.5, along with other clustering methods used in the study.

3.3.1.2 Multi-Resolution Modeling

The next step is to identify the path-level demands on alternative routes associated with each representative scenario. MRM, combined with DTA, provides a way to model the scenarios and estimate the path-level demand. The MRM of the study corridor (I-95 corridor) uses a combination of macroscopic modeling (regional planning model), mesoscopic simulation-based DTA, and microscopic simulation. Figure 3-5 shows the flow chart of the MRM modeling steps used in this study.

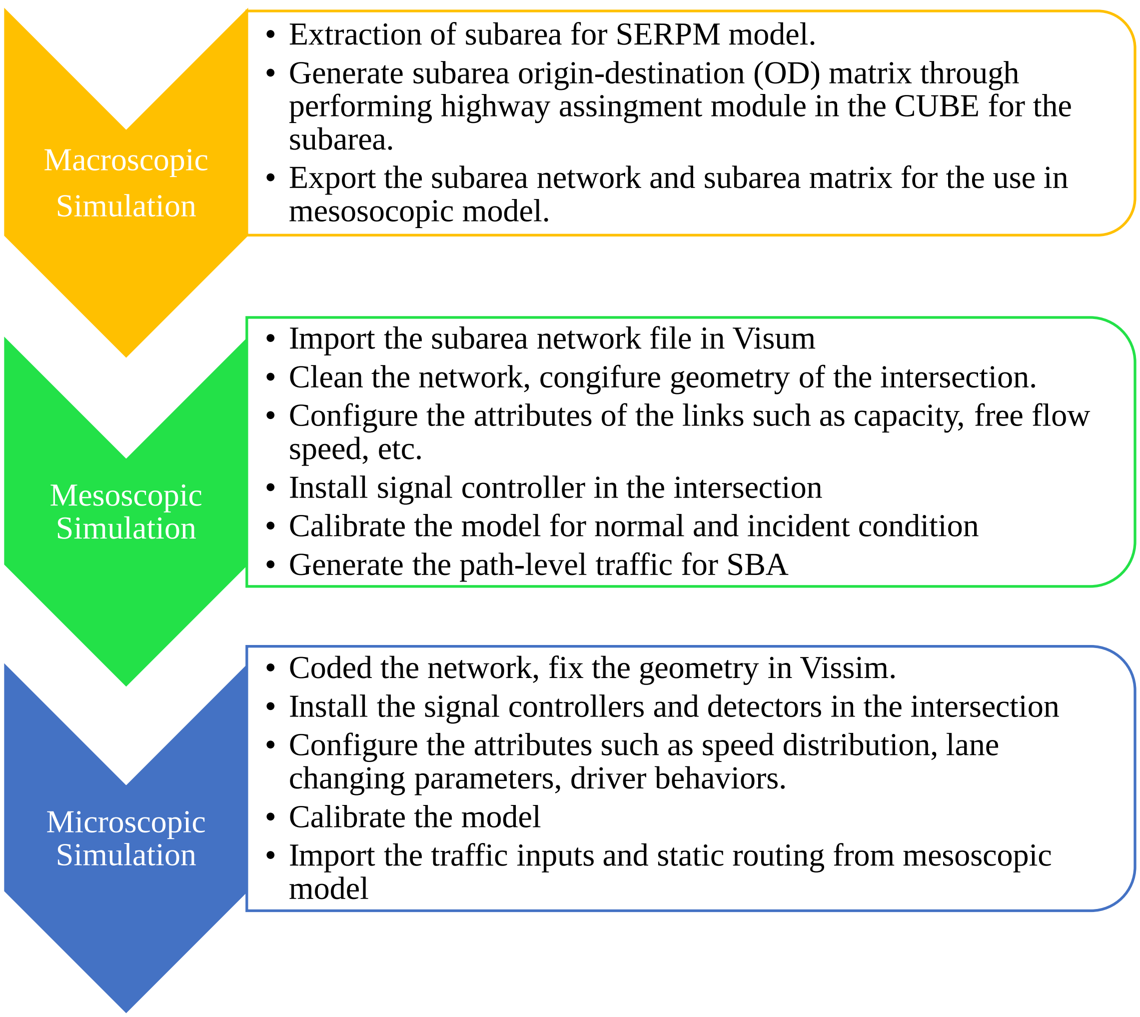


Figure 3-5: MRM Modeling Frameworks and Associated Tasks

Calibration of the Simulation Models

Proper replication of the real world through simulation in different modeling resolutions requires meticulous effort in the calibration of the model parameters so that it can produce traffic volume and travel time in acceptable deviation from the real-world measurement. This study utilizes the 2004 Federal Highway Administration (FHWA) calibration criteria of traffic simulation models, shown in Table 3-1 (FHWA, 2004).

Table 3-1: Calibration Criteria and Acceptance Target for Freeway and Arterials

Criteria and Measures	Calibration Acceptance Targets
Hourly Flows, Model Versus Observed Individual Link Flows: Within 15%, for 700 veh/h < Flow < 2700 veh/h	> 85% of cases
Hourly Flows, Model Versus Observed Individual Link Flows: Within 100 veh/h, for Flow < 700 veh/h	> 85% of cases
Hourly Flows, Model Versus Observed Individual Link Flows: Within 400 veh/h, for Flow > 2700 veh/h	> 85% of cases
Travel Times: Journey Times, Network Within 15% (or 1 min, if higher)	>85% of cases

One of the critical aspects of this study is the calibration of the mesoscopic model for the incident condition. This calibration is performed in two steps. First, the model is calibrated for normal conditions. The calibration for the incident condition then follows this. Figure 3-6 shows the calibration procedure carried out in this study for normal and incident conditions. The link traffic volumes are estimated using the dynamic assignment in the utilized mesoscopic simulation model (Visum) (PTV AG, 2019). The Least Square OD matrix estimation (ODME) module in Visum is used to produce OD demands that provide a good match to the turning movement counts for the normal conditions. The parameters used as inputs to the ODME modules are traffic detector counts retrieved from the FDOT data warehouse and turning movement counts obtained from FDOT District IV.

For the incident condition, the real-world data used in the calibration of the mesoscopic simulation-based DTA model for incident conditions are traffic counts collected by traffic sensors on the freeway mainline, travel time data from a third-party vendor (HERE), and path-specific OD traffic data from another third-party vendor (StreetLight). The path-specific OD traffic data contains an origin zone, a destination zone, and a path that the traffic uses between the origin and the destination. Incident attributes such as the start time, incident duration, number of blocked lanes, and incident location are coded in the model. The capacity reductions due to incident lane blockages are replicated in the model by adjusting the model parameters to meet the recommended reduced capacity due to incidents in the HCM 2016 (HCM, 2016). The process is performed through an iterative process and continued until the calibration criteria are met. After calibration, the resulting path-based OD traffic volumes and travel times are verified against the criteria shown in Table 3-1.

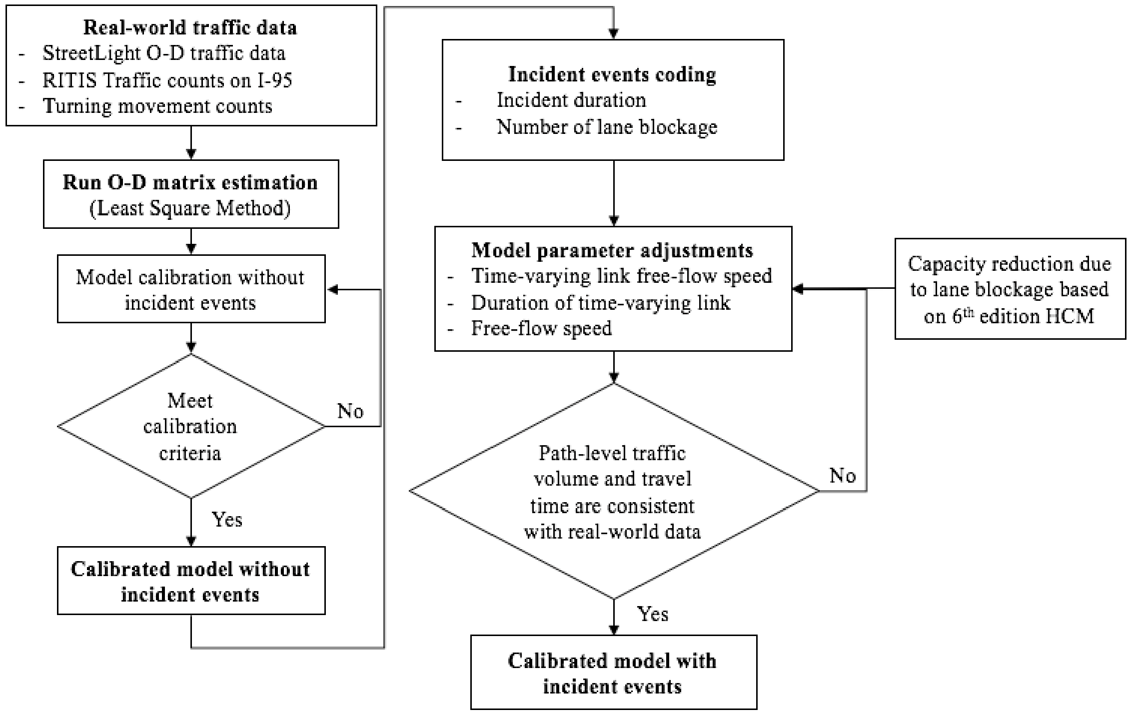


Figure 3-6: Mesoscopic Model Calibration Procedure

A microscopic model in the PTV Vissim simulation tool (PTV AG, 2019) is calibrated following the same criteria set in the FHWA TAT Volume III. The calibrated microscopic model is used to develop and assess signal plans for the identified scenarios.

3.3.2 Optimization of the Signal Plan

One of the variations of the GA, NSGA-II, as identified in the literature review, is used in the study to develop the signal timing plan in this study. NSGA-II is a multi-objective optimization that uses the same basic principle of GA except in selecting individuals for the next generation. In the following two sections, both GA and NSGA-II are discussed in detail.

3.3.2.1 Genetic Algorithm (GA)

The theoretical foundation of GA is originally developed by Holland (1975). It is a heuristic optimization technique that imitates the biological processes of reproduction and natural selection to solve for the 'fittest' solutions (Goldberg and Holland, 1988). It is a stochastic process and far more powerful and efficient than random search and exhaustive search algorithms (Kinnear, 1994).

Features of GA

GA optimization works similar to the biological evaluation process; hence features of the algorithm named after the biological process. Following are the salient features of the GA.

Initialization

GA initiates with a population of individuals; each individual represents a possible solution. The parameters of each individual are encoded into a chromosome. In general, each parameter value is converted into a binary string (1's and 0's) then concatenate the parameters end-to-end like genes in a DNA strand to create the chromosomes (Mitchell, 1995).

Fitness assignment

Each candidate solution's fitness is then evaluated with respect to a given objective function and assigned a fitness value. The higher the fitness value, the better the solution and possess the greater probability of being selected for recombination.

Selection

After assigning the fitness value, the fittest individuals are selected for the crossover to produce the next generation. Elitism selection makes the fittest individuals survive directly for the next generation.

Crossover

The highly fit individuals selected in the selection process are given opportunities to reproduce by exchanging their genetic information in this step. Each offspring of the new population represents a new solution, which shares some of the characteristics obtained from both parents.

Mutation

Some genes in the string of each individual are altered through the mutation process. By this process, offspring can replace the whole population or less fit individuals. The mutation is a very crucial step in the optimization process. Although selection and crossover keep the fittest individual, these are only fitter relative to the current population. This can cause the algorithm to converge too quickly and lose "potentially useful genetic material (1's or 0's at particular locations)" (Goldberg and Holland, 1988). In other words, the algorithm can get stuck at a local optimum before finding the global optimum (Haupt and Haupt, 2004). The mutation operator helps protect against this problem by maintaining diversity in the population, but it can also make the algorithm converge more slowly.

3.3.2.2 Non-dominated Sorting Genetic Algorithm-II (NSGA-II)

NSGA-II belongs to a set of multi-objective algorithms that strive to find the Pareto front of compromised solutions of all objectives rather than integrating all objectives together (Deb et al., 2002). The algorithm works similarly to the basic GA with the crossover and mutation steps. A selection operator is used to create a mating pool by combining the parent and offspring populations and selecting the best individuals following the process of the non-dominated sorting and crowding distance sorting.

Non-dominated sorting is based on the concept of non-dominated sets. A solution *'x'* is said to dominate a solution *'y'* if *'x'* is no worse than *'y'* in all objectives and if *'x'* is strictly better than *'y'* in at least one objective. For a set of solutions P, the non-dominated set is formed by those solutions that are not dominated by any member of P. The goal in multi-objective optimization is to find the set of Pareto-optimal solutions, i.e., the non-dominated set of the entire feasible search space.

To calculate the crowding distance, the distances between the left and right neighbors of a solution are summed up for all objectives. Solutions that are located in a region of the search space that is not densely covered by the current population are preferred. The steps of the NSGA-II are described below.

Steps

i. A random parent population P_t is initially generated. The population is sorted based on the non-domination. Each solution is assigned a fitness (or rank) equal to its non-domination level. The usual binary tournament selection, crossover, and mutation operators are used to create an offspring population Q_t of size N.

ii. A combined population $R_t = P_t \; U \; Q_t$ is formed. The population R_t is of size $2N$. Then, the population R_t is sorted according to non-domination and assigned a rank to each population.

iii. The population R_t is again ranked based on their crowding distance.

iv. The next generation parents' P_{t+1} are selected in a way that is, between two individuals (solution) with differing non-domination ranks; the lower rank individual is selected. But if both individuals belong to the same front, then the individual with higher crowding distance is selected.

v. The processes loop through Step-ii until a termination condition meets.

3.3.3 Pareto Front

Francis Ysidro originally introduces the Pareto optimality concept and then generalized it by Vilfredo Pareto (Coello et al., 2007). A solution belongs to the Pareto set if there is no other solution that can improve at least one of the objectives without degradation of any other objective. The Pareto set selection is explained using a hypothetical example shown in Figure 3-7, where two objectives' functions (f_1 and f_2) are minimized. Among the set of feasible solutions, *A* and *B* are included in the Pareto front. At the same time, *C* and *D* are left out from the Pareto front because solution *C* does not improve any of the objective functions, and *D*, although it improves f_1 but degrades the f_2.

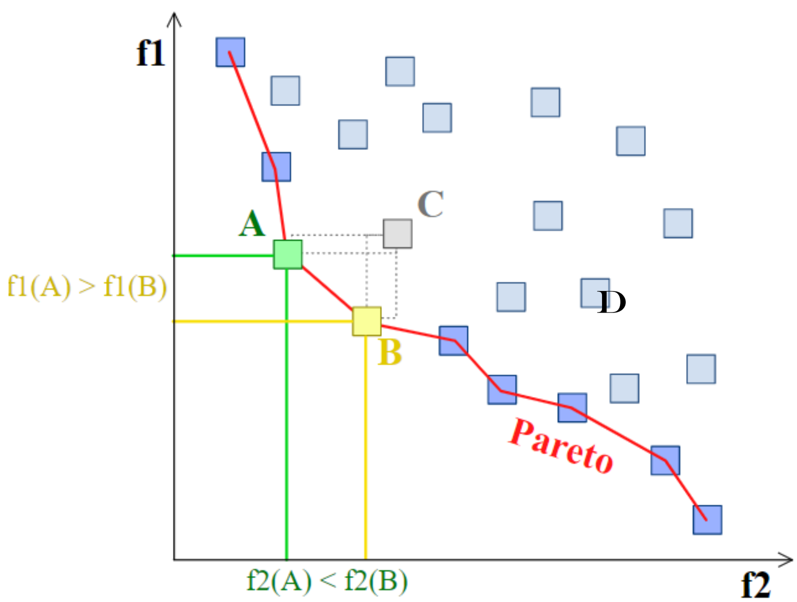


Figure 3-7: Pareto Front for Two Objectives Minimization Problem (Source: Nojhan, 2021)

3.3.4 Optimization Problem Formulation

A microscopic simulation-based multi-objective optimization is utilized in this study to develop the signal plans corresponding to different diversion scenarios during incidents on the freeway. The Vissim microscopic model developed as a part of MRM modeling is used to simulate the scenarios, allowing the estimation of the objective functions' values for the feasible solutions. The NSGA-II algorithm is used to solve the optimization problem. Since diversion can create congestion on intersection movement(s) with long queues, the objective functions used in the optimization include throughput maximization of the impacted movements and overall delay minimization of all the intersections movements along the diverted path. The objective functions and subjected constraints used in the optimization are the following.

$$\text{Minimize } f_1(g) = \left(\sum_{i \in I} \sum_{m \in M} V_{i,m} \times d_{i,m}\right) / \sum_{i \in I} \sum_{m \in M} V_{i,m} \tag{3-10}$$

$$\text{Maximize } f_2(g) = \sum_{i \in I} \sum_{m \in M} N_{i,m} \tag{3-11}$$

The objective functions are subjected to the following constraints.

$$Cmin_i < C_i < Cmax_i \qquad \forall i \in I \tag{3-12}$$

$$gmin_{k,i} < g_{k,i} < gmax_{k,i} \qquad \forall i \in I, \forall k \in K \tag{3-13}$$

$$g_{1,i} + g_{2,i} = g_{5,i} + g_{6,i} \qquad \forall i \in I \tag{3-14}$$

$$g_{3,i} + g_{4,i} = g_{7,i} + g_{8,i} \qquad \forall i \in I \tag{3-15}$$

where

$f_1(g)$ = average delay,

$f_2(g)$= total throughput,

$d_{i,m}$= average delay for movement m, at intersection i,

$V_{i,m}$ = number of vehicles for movement m, at intersection i,

$N_{i,m}$ = total throughput for movement m, at intersection i,

C_i= cycle length of intersection i,

$Cmin_i$= minimum cycle length of intersection i,

$Cmax_{\mathrm{i}}$= maximum cycle length of intersection i,

$g_{k,i}$ = green duration for phase k, at the intersection i (decision variables),

$gmin_{k,i}$ = minimum green time associated with phase k, at the intersection i,

$gmax_{k,i}$= maximum green time associated with phase k, at the intersection i,

$g_{1,i}$= northbound left phase split at the intersection i,

$g_{2,i}$ = southbound through phase split at the intersection i,

$g_{3,i}$ = eastbound left phase split at the intersection i,

$g_{4,i}$ = westbound through phase split at the intersection i,

$g_{5,i}$ = southbound Left phase split at the intersection i,

$g_{6,i}$ = northbound Through phase split at the intersection i,

$g_{7,i}$ = westbound Left phase split at the intersection i,

$g_{8,i}$ = eastbound Through phase split at the intersection i,

I = set of all intersections of the alternative route,

M = set of all movements at the intersection i, and

K = set of all phases available at the intersection i.

The constraints used in the optimization are the cycle length, minimum and maximum green times, and the ring and barrier settings. The constraints are shown in Equations 3-12 to 3-15. The cycle length constraint in Equation 3-12 shows a bounded value between a minimum value and a maximum value set in the optimization. However, in the case study used in this dissertation, the cycle length constraint is further modified to Equation 3-16.

$$C_i = Cexist_i \quad \forall i \in I \tag{3-16}$$

where,

$Cexist_i$ = existing cycle length at intersection i.

Equation 3-16 indicates that, in the case study, the cycle lengths for all intersections in the alternative routes are kept equal to the existing cycle lengths used in the field to ensure no violation of progression of the whole alternative route, since only a part of the alternative route is modeled in this study.

The minimum green time and maximum green time constraints are shown in Equation 3-13. In calculating the minimum and maximum green times for each phase for all intersections, this study utilizes the existing minimum and maximum green times used in the field. These values are utilized to maintain the pedestrian crossing time requirements. The barrier constraints are used to avoid the operation of conflicting movements at the same time. The barrier separates the north-south movements from the east-west movements. Equations 3-14 and 3-15 show the constraints that reflect the ring and barrier settings of the controller.

In the optimization, the decision variables are the green splits of all phases of the intersections in the diversion routes. The green splits are optimized based on the values of objective functions estimated from the simulation. For different green split values, the simulation provides different values of objective functions. The following two sections describe the population design and optimization process.

3.3.5 Population Design and Parameters Configuration

All the intersections of the study area are four-leg intersections, operating following the NEMA 8-phase signal control. A typical signal control phase for a four-leg intersection is shown in Figure 3-8. It consists of lead, lag, and overlap phases.

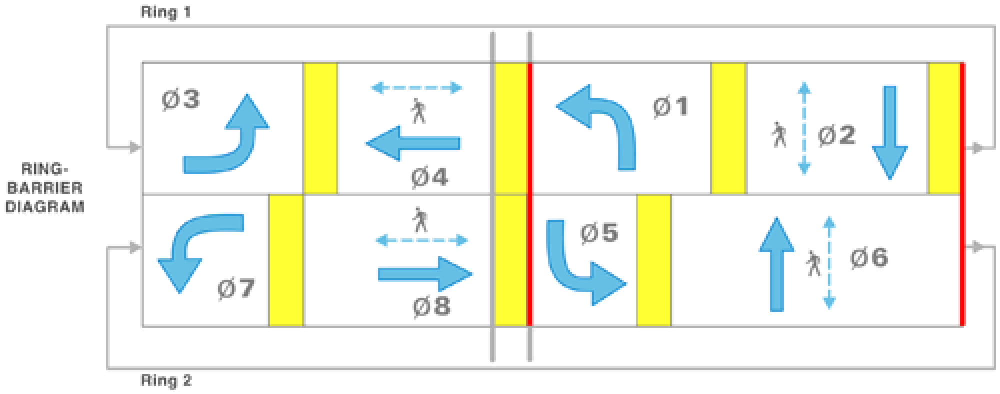


Figure 3-8: Typical Signal Control Plan (Source: FHWA, 2021)

To encode the phases in the candidate solution (i.e., chromosome), all the phases are coded in the binary 0's and 1's and concatenated with each other. The MRM study network consists of three intersections at the microscopic simulation level, for which the signals are optimized. All eight-phase splits for all three intersections were encoded together in a chromosome. Each candidate solution (aka chromosome) length is 192 bits (i.e., genes), and the population size used in the study is 20. A total of 25 generations is used in the study to find the optimal Pareto front. In a Vissim based GA optimization, Stevanovic et al. (2007) found that the best crossover probability between 0.4 and 0.7 and mutation probability between and 0.01 and 0.04. Therefore, in this analysis, the utilized crossover rate is 0.6, and the utilized mutation rate is 0.03. The yellow and all-red intervals are added to the corresponding phase splits.

3.3.6 Optimization Process

The microscopic simulation-based optimization is performed using the Vissim component object model (COM) service, an interface that provides communication between software (Box, 1998). The Python programming language is used to call both the COM and NSGA-II algorithm. The COM interface is used to run the simulation and to access, modify, and retrieve different attributes of the simulation model. The attributes retrieved from the model are used as inputs to the NSGA-II optimization, and the output is fed into the simulation model. After developing the communication between the simulation model and the COM, a COM interface calls the NSGA-II to generate the initial population, which is a set of individuals consisting of genes, representing the green split of the phases. In generating the initial population, the green splits and the cycle lengths of all signal plans of all intersections are checked using an algorithm to make sure they fulfill the constraints of the optimization. In the next step, the simulation is run by the COM, and the current step green splits of the traffic controller are changed according to the individual generated in the previous step. At the end of each run, the objective functions' values are estimated using simulation for each individual. The offspring solutions are then generated following crossover and mutation operation. In this step, the constraints are also checked for all of the offspring solutions. Then, the simulation is run for all of the offspring solutions, and the outputs are saved. After assigning the outputs to each individual, the non-dominated sorting and crowding distance sorting are estimated, and the population for the next generation is selected.

For each individual (a signal control plan), the simulation is run for 30 minutes (1,800 seconds) where the first 15 minutes are considered the warm-up period and are not used in the performance measurement. The last 15 minutes' (900 seconds) outputs are saved and averaged across the cycles to use as the objective functions' values. The pseudo-code of the algorithm in the Vissim platform is shown in Figure 3-9.

NSGA-II Algorithm with N Generations and 2 Objective Functions

Input: P_0, N, M
P_0= Initial Population
M= Population size
N= Number of Generations
Output: Pareto Solutions

Initialize
N=0;
Generate Initial Population, P_0
Run Vissim simulation
Break Vissim simulation at 900 seconds.
Run Vissim simulation using COM interface for individuals in P_0
Compute fitness f_1, f_2 using Vissim COM interface
while termination conditions are not satisfied ($N<N_{max}$) do:
    Convert individual of P_N into a binary variable
    Q_N= Crossover & Mutation (P_N)
    Reconvert individual of Q_N into decimal variable
    Run Vissim simulation
    Break Vissim simulation at 900 seconds.
    Run Vissim simulation using COM interface for individuals in Q_N
    Compute fitness f_1, f_2 using Vissim COM interface
    $R_N = P_N \cup Q_N$ (size 2M)
    P_{N+1}= [ ]
    (L_1, L_2, …, L_i) = Ranked R_N based on Non-dominated sorting and Crowding distance
    i=1
    repeat
        $P_{N+1} = P_{N+1} \cup L_i$
        i=i+1;
    until $|P_{N+1}| = M$
    if $N=N_{max}$ **then**
    break
    else, N=N+1
end
return Pareto-optimal front and associated signal timing plan.

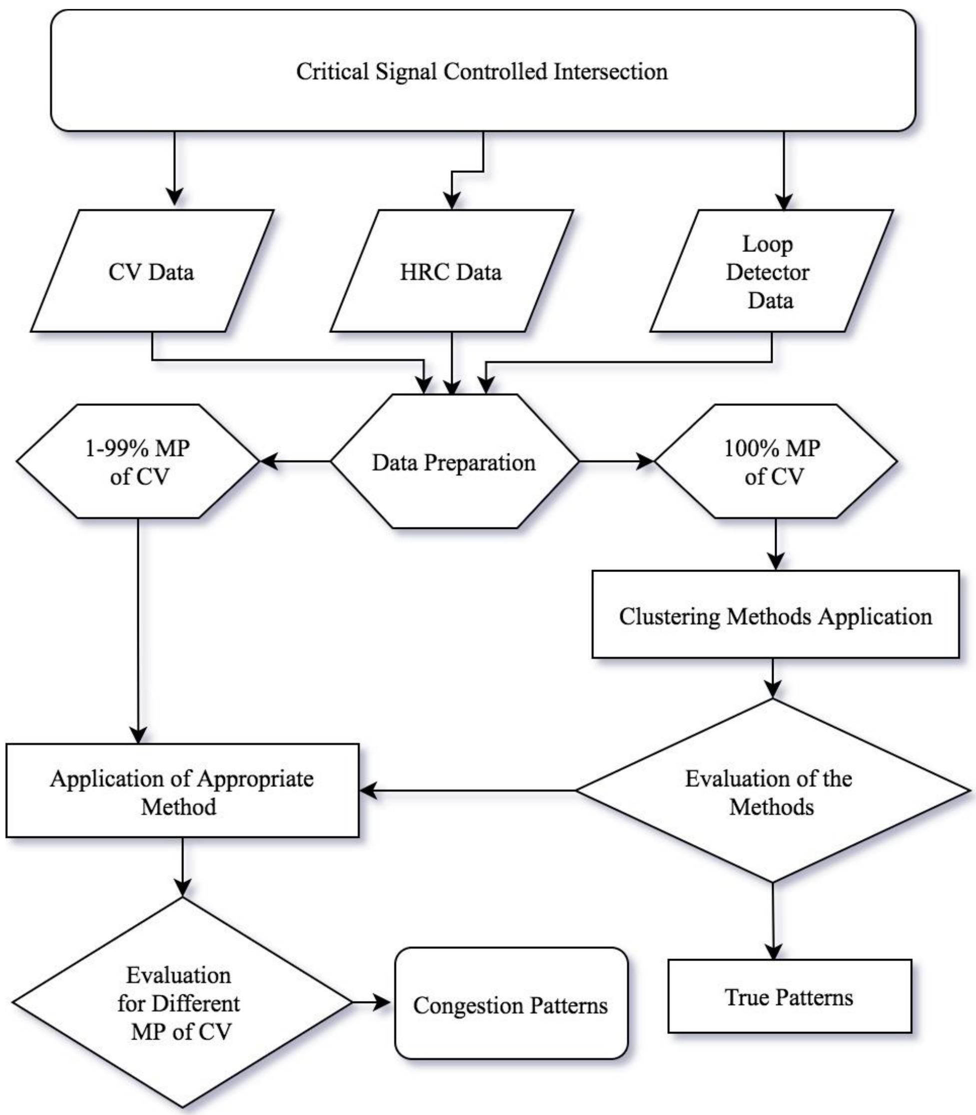
Figure 3-9: NSGA-II Algorithm Pseudo Code

3.4 Implementation of the Signal Plans in Real-Time

The methodology described in this dissertation derives the signal timing plans off-line utilizing optimization methods. These plans can then be stored in a library of special signal timing plans for activation in real-time. In addition, the off-line data analytics utilized in this study identify the conditions under which each of the plans should be activated. The agency can use these methods that are done off-line based on archived data to select the signal plans on the critical alternative routes in real-time from the library of plans based on the measured travel times and incident status. The critical diversion routes can be predicted based on the *ΔTT value* estimated based on real-time measurements utilizing the developed LSTM models. At the same time, this *ΔTT* value can also be used to predict the operation scenario in the next 90 minutes and select the signal plan developed for that scenario. After every 15 minutes, the travel time and incident status can be updated in the LSTM models, and the same procedure to implement the signal plans is repeated. The process continues till the predicted *ΔTT* values fall below the set threshold values. The selection and implementation of the signal timing plans in real-time operation from a library of plans developed off-line, is recommended to be implemented as part of the central software of traffic management centers.

3.5 Congestion Identification using CV and HRC

One of the crucial aspects of the study is the development of the signal timing plans for the critical intersections of the alternative routes based on the predicted traffic scenarios that originated from the diversion. These scenarios can vary extensively based on the incident and traffic attributes both on the freeway and the alternative routes. Moreover, an important consideration is the distributions of the diverted traffics at the critical intersections. For example, sometimes a major portion of the diverted traffic takes a left turn to a parallel arterial street at an intersection in the close vicinity of the freeway, while sometimes the diverted traffic goes through at the intersection to use another parallel route further away from the freeway. This asymmetric distribution of the traffic at the intersection affects congestion by creating long queues, queue spillback, and queue spillover. The MRM method described in the previous section provides an off-line simulation-based solution in identifying these scenarios effectively and developing signal timing plans for those scenarios. Because the traditional data collection techniques such as stop line detector or the Bluetooth detector sometimes cannot discern between the contributions of different movement groups of an approach to congestion. These existing technologies may not be able to capture, for example, if the high delay and queue on an approach are due to the left turn or through movements, which is important to the selection of the best signal plans. Therefore, utilizing data from existing technology without additional data in designing the signal plan may lead to the wrong assignment of green time to movements. Contrary to these traditional detectors, CV and HRC are the two new imminent technologies that have shown a lot of promise in traffic management. This study explores the use of these data in the identification of congestion patterns on alternative routes. These patterns can be used to select the signal plans in real-time. A clustering-based approach is developed to identify the congestion patterns based on the data obtained from CV and HRC. The process followed in identifying the congestion patterns is shown in Figure 3-10. Initially, the CV, HRC, and loop detector data are merged, and critical attributes such as average travel time, CV position in the queue, served volume and capacity ratio, green occupancy ratio, red occupancy ratio, mid-block volume are measured. These attributes are taken for 100 percent market penetration (MP) of CV and for other percentages as well. The 100% MP CV data is used for determining the appropriate clustering method and true patterns as no real-world traffic patterns are available in the study location. Finally, congestion patterns for various MP of CV are measured and compared against the true patterns to assess the accuracy.



CV=Connected Vehicle, HRC=High-Resolution Controller, MP=Market Penetration

Figure 3-10: Steps of Congestion Patterns Identification using HRC and CV Data

As mentioned previously, the study area does not have real-world data of CV or HRC. As a result, the study utilized Vissim microscopic simulation to simulate all the data. The detailed data generation process is described in Chapter IV.

3.5.1. Clustering

Clustering is the best available unsupervised machine learning technique for identifying the patterns in the data (Saha et al., 2020; Nafis et al., 2021). Four clustering techniques, two linear and two non-linear clustering methods are explored, and their performances in the congestion identification are evaluated in the study. Below is the description of the clustering techniques that are tested for use in this study.

3.5.1.1 K-means

The K-means algorithm is a widely used method applicable for clustering data based on quantitative variables (Jain and Dubes, 1988). The method is based on an iterative algorithm in which the process is initiated by providing a fixed set of centroids (Hartigan and Wong, 1979). Each data point to be clustered is then assigned to its closest centroid using a squared Euclidian distance measure. To assign a point to a cluster, the goal is to minimize the sum of average pair-wise distance within-cluster dissimilarity. The centroids are then updated by computing the average of all the points assigned to each cluster. The steps are iterated until the assignment of the data points to each centroid does not change significantly.

3.5.1.2 Principal Component Analysis (PCA) Combined with Clustering

Principal Component Analysis **(**PCA) is a statistical approach for dimension reduction and compression while retaining most of the variation in the data set (Dunteman, 1989). PCA converts the observations to an orthogonal system of Euclidean space and thus reduces dimensionality by retaining only those characteristics of the data set that contribute most of its variance. The first principal component (PC) can equivalently be defined as a direction that maximizes the variance of the projected data. The i^{th} PC can be taken as a direction orthogonal to the first i^{th}-1 PCs that maximizes the variance of the projected data. A small number of PCs from all PCs that preserve optimum variability are used for further analysis. PCA was found effective in capturing the cluster structure in the data set when used along with clustering methods instead of clustering methods by themselves (Meng et al., 2015). Ding and He (2004) found that K-means clustering on high dimension data was affected by the noise in the data set, and applying K-means clustering in the PCA subspace improved the results significantly. Clustering with the reduced dimensions from PCA was found very effective in recognizing the patterns in the data set (Saha et al., 2019).

3.5.1.3 t-Distributed Stochastic Neighbor Embedding (t-SNE)

Unlike the traditional dimensionality reduction techniques such as PCA, t-distributed stochastic neighbor embedding (t-SNE) is a non-linear dimension reduction technique that keeps dissimilar data points far away in low dimensional representations. t-SNE is a variation of stochastic neighbor embedding (SNE), which uses Student-t distribution rather than a Gaussian to compute the similarity between two points in the low-dimensional space (Maaten and Hinton, 2008). The method uses equality of conditional probabilities representing similarities between the data points with high-dimension and low dimension based on the Euclidean distances in the dimension reduction. The sum of Kullback-Leibler (KL) divergence is minimized over all data points based on a gradient descent method. Although SNE constructs reasonably good visualizations, it is hampered by a crowding problem that makes cost function very difficult to optimize. t-SNE employs a heavy-tailed Student t-distribution in the low-dimensional space to alleviate both the crowding problem and the optimization problems of SNE (Maaten and Hinton, 2008). t-SNE has been shown to successfully identify small cellular subpopulations, as low as those comprising 0.25% (Amir et al., 2013).

3.5.1.4 Deep Embedded Clustering (DEC)

Deep Embedded Clustering (DEC), which learns feature representations for clustering tasks using deep neural networks, has attracted increasing attention for various clustering applications. The method simultaneously learns feature representations and cluster assignments using deep neural networks. It is a parameterized non-linear mapping from the data space X to a lower-dimensional feature space Z through the optimization of stochastic gradient descent (SGD) via backpropagation. It is a method that simultaneously solves for cluster assignment and the underlying feature representation (Xie et al., 2016). Deep neural networks training in this method is different from other supervised learning, as there is no labeled data associated with the analysis. The method refines clusters iteratively with an auxiliary target distribution derived from the current soft cluster assignment. This process gradually improves the clustering as well as the feature representation. DEC was applied successfully in various filed such as prediction of severity of age-related macular degeneration (AMD) from input optical coherence tomography (OCT) images (Mahapatra et al., 2020), and pattern detection from seating pressure distribution during wheelchair motion (Noguchi et al., 2019).

3.5.2 Optimum Number of Clusters

One of the important aspects of clustering is to determine an adequate number of clusters to discern all the frequent patterns. Several empirical methods are available to identify the required number of clusters, such as the Elbow Method, Average Silhouette Method, and Gap Statistics Method. In this study, the optimal number is selected using the Elbow Method. With this method, a graph is drawn between the sum of square error (SSE) measure and the number of clusters, and the location of the bend in the plot is used as an indicator of the appropriate number of clusters (Ketchen and Shook, 1996).

Besides the Elbow method, the study also utilizes the t-distributed stochastic neighbor embedding (t-SNE) method (Maaten and Hinton, 2008), a data visualization technique to verify the optimum number of clusters determined in the Elbow method. t-SNE reduced high-dimension data to two dimensions data that makes it easily visualizable in the 2-D graph. The visualization graph depicts the inherent pattern of the data and provides a prediction of the optimum number of clusters.

3.5.3 Evaluation of Clustering Methods

The performance of the investigated clustering methods is assessed utilizing the Silhouette Coefficient (Rousseeuw, 1987) and Davies-Bouldin Index (DBI) (Davies and Bouldin, 1979). These two performance measures are chosen in the study for their capability to assess the performance of the clustering algorithms. These measures do not need ground truth data and allow a straightforward interpretation of the results (Rousseeuw, 1987). A higher Silhouette coefficient indicates a dense and well-separated cluster, while a lower DBI describes a higher degree of appropriateness of data partitions into clusters (Davies and Bouldin, 1979). Based on the coefficients’ values, the best clustering method is selected to identify the congestion patterns.

3.5.4 Analysis of Congestion Patterns

The congestion patterns in this study are initially identified considering the presence of 100% CV vehicle in the traffic stream for all the possible diversion scenarios. This is done as the study network doesn't have any ground truth data for congestion patterns. After identifying the congestions for the base scenario, this study measures the congestions for other market penetrations (MP) of CV and evaluated the accuracy against the base scenario.

3.6 Summary

This research focuses on developing methods and associated models to support the management of diversion routes utilizing a combination of data analytics, machine learning, and simulation modeling. In the first step, a methodology is developed utilizing machine learning to identify the critical routes used by the motorist during incidents on the freeway. In the next step, a methodology consisting of clustering analysis, MRM modeling, and multi-objective optimization is developed to generate special signal plans for the critical routes. Finally, a methodology is developed to explore the use of CV and HRC data for identifying traffic congestion patterns, which can be used to select the signal timing plans.

CHAPTER IV

CASE STUDY

This chapter describes the case study and the associated data (i.e., sources, retrieval, and preparation) utilized in this research to demonstrate the methodology discussed in Chapter III. The first section explains the network selection and the rationale of choice for the study. The second section describes the network preparation process and its salient features. Each model developed in the study requires a specific format of the data; the third section describes the retrieval and formation process of the data.

4.1 Network Selection

The networks are selected based on criteria such as the level of congestion created by the incident, availability of data, and the potential for improvement by introducing special signal timing plans on alternative routes used by diverted traffic. The utilized network is the network surrounding Interstate-95 (I-95) in Broward and West Palm Beach County, FL. The I-95 is one of Florida's busiest corridors, with annual average daily traffic (AADT) of over 328,000 in Broward County (FHWA, 2017). The FDOT has deployed ITS based technologies to improve the corridor performance that generates a lot of data. Besides, this I-95 segment is also considered a potential site for ICM implementation by the FDOT and other agencies in the region.

The I-95 corridor is utilized to demonstrate the methodology described in Chapter III for critical route prediction, MRM modeling to estimate route demands, and special signal timing plan generation for the diversion routes. A portion of the same network is also used to identify congestion patterns in the presence of CV and HRC data. Figure 4-1 shows the network in the Open Street Map (OSM) platform.

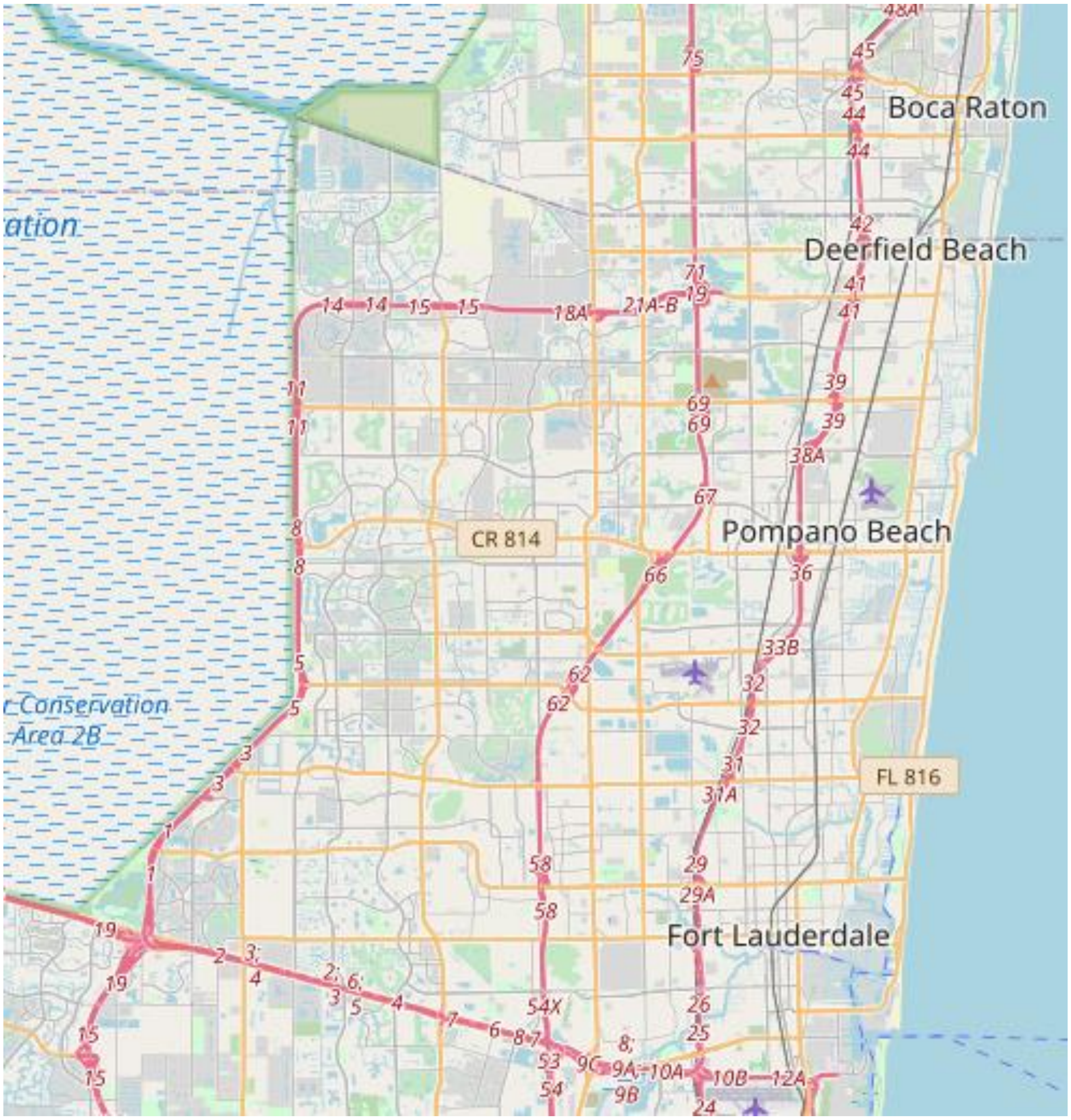


Figure 4-1: I-95 Corridor Traffic Network

4.2 Network Preparation

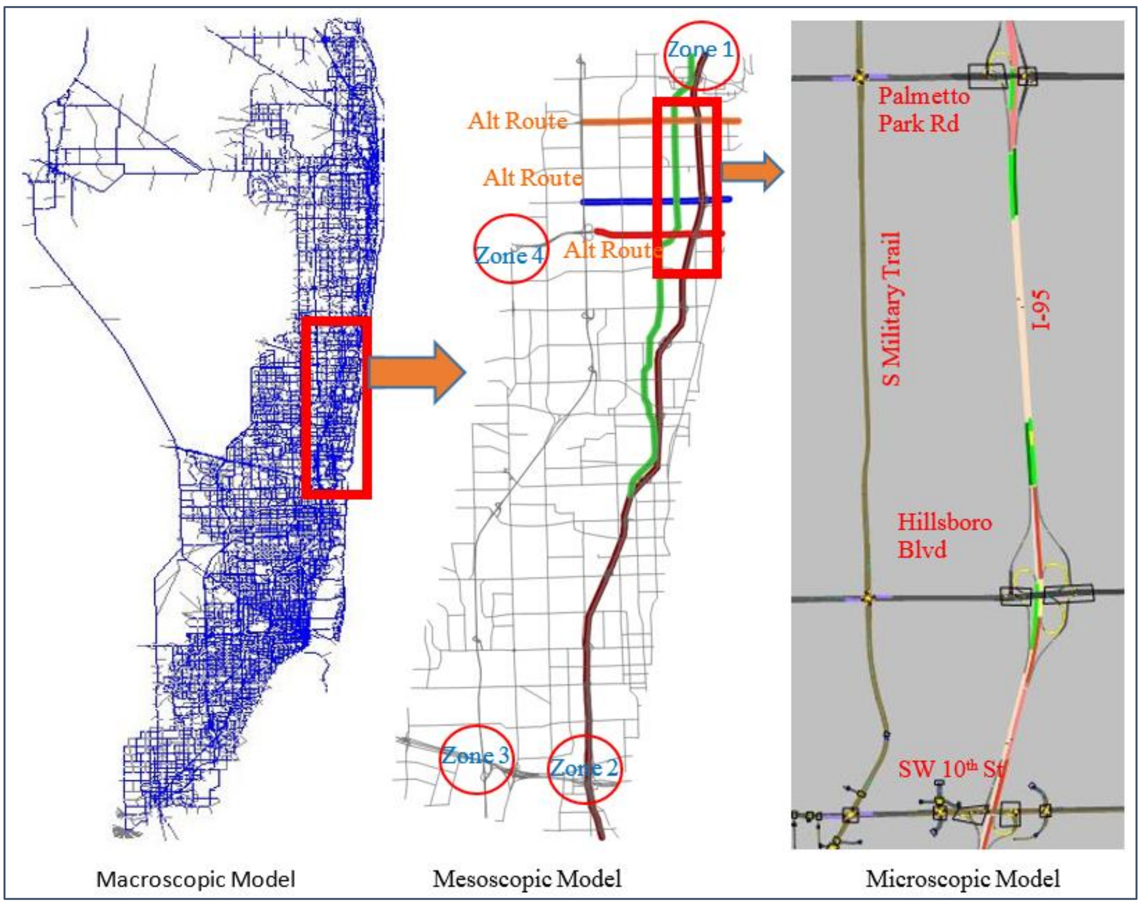
After selecting the networks, the networks are coded in the simulation platform to develop and evaluate the models. Three different resolutions: microscopic, mesoscopic, and microscopic of the I-95 corridor, are prepared using Cube Voyager from Citilabs and two PTV traffic simulation software suits – Visum and Vissim, respectively. All three resolutions of the MRM network are shown in Figure 4-2, as described below.

I. Cube Voyager is a macroscopic simulation tool and currently in use of the Southeast Florida Regional Planning Model (SERPM) that encompass the I-95 corridor. The I-95 corridor subarea network is extracted from SERPM along with the OD matrix for the subarea network. The highway assignment module of the Cube Voyager is applied to extract the OD matrix for a total of 435 zones, including internal and external zones.

II. The mesoscopic simulation is performed using Visum software. Both the subarea network and OD matrix generated in the previous step are imported in Visum through the "VisumAddIn" service. After importation, the capacity of the links, intersection geometry, and movements are rectified in Visum to match the real world. The mesoscopic simulation is performed using simulation-based assignment (SBA) that considers traffic signal control. The traffic signal controls in the majority of the critical links surrounding the I-95 are set for proper SBA assignment. The OD demands used in the assignment are obtained using an ODME process.

III. The microscopic simulation model, coded using Vissim, consists of both arterial and freeway segments. The freeway segment, I-95, extends from beyond the off-ramp of Glades Road to the off-ramp of West Atlantic Blvd. Parallel to I-95, the South Military Trail arterial is considered as one of the potential diversion routes and therefore was coded in the network. Three important connectors between I-95 and South Military Trail that facilitate the diversion operation for the South Bound (SB) traffic are also coded in the network. The connectors are W Palmetto Park Rd, W Hillsboro Blvd, and SW 10th St., whose lengths vary in a range of

0.5-0.6 mile. Signal control plans of these connectors' intersections are then coded as actuated controlled signals based on the obtained data from Broward County. This Vissim network is also used to develop and evaluate the signal control plan for different diversion scenarios. The S Military Trail – W Palmetto Park Rd intersection of this network is used for the congestion pattern identification using CV and HRC data.



(a) SERPM Modeled Area **(b) Subarea** **(c) I-95 and major arterial**

Figure 4-2: Networks for the Three Levels of Modeling

4.3 Data Retrieval

Two types of data are mainly utilized in the study; i) real-world data and ii) simulated data. Real-world data includes travel time and traffic data for I-95, incident data for I-95, travel time for arterials, origin-destination (OD) matrix data from StreetLight. The simulated data includes CV data and HRC data for the I-95 corridor. Simulation data are used in the study as the study area does not currently have these data. Besides, simulation provides the flexibility to change the demands to examine the impact of demands on the analysis results.

4.3.1 Traffic Data

The analysis period is set from January 1st, 2017, to December 31st, 2018, excluding holidays and weekends. Traffic data, including volume and speed measurements, are collected from point sensors along I-95 mainline. The traffic data is retrieved from the FDOT data warehouse, which is a part of the Regional Integrated Transportation Information System (RITIS) in 15-minute intervals for four periods (i.e., AM Peak (7:00 AM – 9:00 AM); Midday (9:00 AM-4:00 PM); PM Peak (4:00 PM-7:00 PM); Evening (7:00 PM-10:00 PM)) in the southbound direction of I-95.

4.3.2. Incident Data

Incident data for the analysis horizon is retrieved from the incident management database maintained by the Florida Department of Transportation (FDOT) District IV. The data includes detailed attributes of the incidents used in the analysis, including incident start time, incident duration, total incident clearance time, number of blocked lanes, severity, and location.

4.3.3 Travel Time Data

The travel time data for both the freeway and alternative routes are retrieved from HERE (a private-sector travel time data provider) for the analysis horizon. One-minute resolution data are obtained for the southbound direction of I-95, the westbound direction of the arterial connectors to the alternative routes, and the parallel southbound arterials. Figure 4-3 shows the I-95, connectors, and parallel arterial for which all these data are collected.

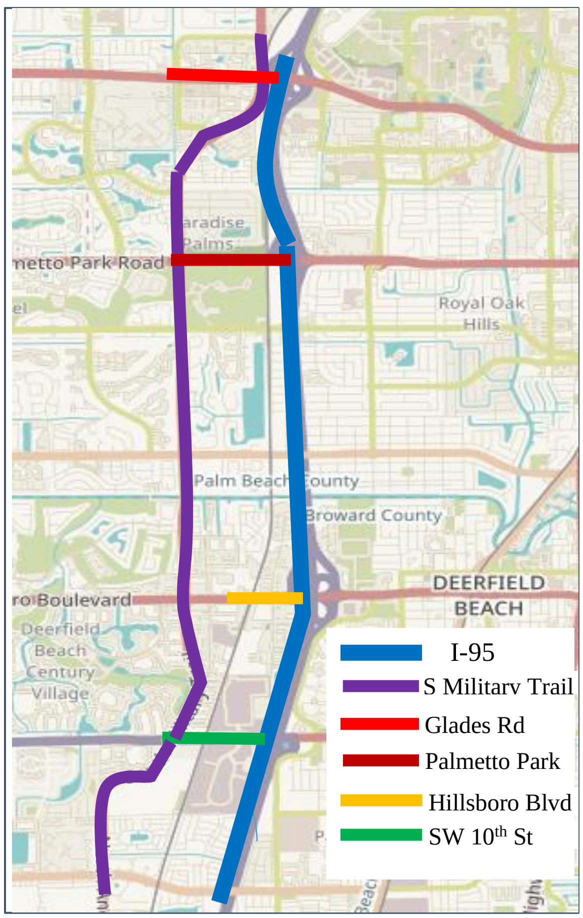


Figure 4-3: I-95 Corridor Diversion Route

4.3.4 StreetLight Data

StreetLight is a third-party vendor that provides path-based travel data of the traffic network. It measures the diverse travel patterns of the traffic from origin to destination using a machine learning algorithm. StreetLight provides both the origin-destination matrix and path-based traffic data. The path-based OD traffic data contains an origin zone, a destination zone, and a path that the traffic uses to reach the destination. The data are available in hourly and AADT format. The study utilized global OD and path-based hourly OD data for both normal and incident conditions. The global OD represents the OD matrix without the path information. The global OD matrix data is used to verify the SERPM OD matrix data. For the calibration of the network during the incident condition, the study mainly utilized path-based traffic data.

4.3.5 Emulated CV Data

The CV data used in the study is emulated using simulation as the network does not have any real-world CV data. Vissim simulation is used to generate the vehicle trajectory files for possible diversion scenarios in the routes. The scenarios are generated for the diversion estimated using the method developed by Tariq et al. (2019). Based on the study, the maximum amount of diversion is 25% of the mainline volume. Therefore, the simulation is run for each percentage of diversion up to 25% by varying the proportion of diverted traffic in the left and through movement. Moreover, the simulation model is also run for different seed numbers.

The vehicle trajectory file, retrieved from the Vissim runs for the scenarios, is a comma-delimited file (CSV) that provides all the vehicle's trajectory information. The file contains vehicle ID, time in seconds from the beginning of the trajectory, the speed of the vehicle in mph, and the x and y coordinates of the vehicle location in feet. For emulating the CV data from the trajectory, a certain percentage of randomly selected vehicles are considered CV vehicles. Their information is used to estimate the parameters used in the analysis.

4.3.6 Emulated HRC Data

HRC data utilized in this study are emulated using Vissim microscopic simulation. The real-world HRC data consist of the signal output states, and the presence of traffic as measured by detectors at 0.1-second intervals is automatically stored in the controller. The data is later retrieved from the controller using a transmission control protocol/internet protocol network connection, with the controller serving as a file transfer protocol (FTP) server. Once downloaded from the controller, a Windows-based wizard is used to convert the binary data files into comma-separated-value (CSV) files for use in producing the appropriate reports (Smaglik et al., 2007).

A sample of real-world controller data is shown in Table 4-1. The data consist of three columns: “Timestamp,” “Event Type,” and “Parameter.”

Table 4-1: High-Resolution Controller Data

Controller No.	Time Stamp	Event Type	Parameter
CAF5C7E2-E505-45B8-8744-ABC5A750F233	11/6/19 7:00:00	44	1
CAF5C7E2-E505-45B8-8744-ABC5A750F233	11/6/19 7:00:02	81	46
CAF5C7E2-E505-45B8-8744-ABC5A750F233	11/6/19 7:00:03	81	47
CAF5C7E2-E505-45B8-8744-ABC5A750F233	11/6/19 7:00:05	22	6
CAF5C7E2-E505-45B8-8744-ABC5A750F233	11/6/19 7:00:07	7	1
CAF5C7E2-E505-45B8-8744-ABC5A750F233	11/6/19 7:00:08	8	1

Sturdevant et al. (2012) defines the enumerations used to encode events on traffic signal controllers with high-resolution data loggers. Table 4-2 shows an example of some of the enumerations.

Table 4-2: Example of Enumeration of High-Resolution Data (Source: Sturdevant et al., 2012)

Event Code	Event Description	Parameter	Description
Active Phase Events:			
0	Phase On	Phase # (1-16)	Set when NEMA phase on becomes active, either upon the start of green or walk interval, whichever occurs first.
1	Phase Begin Green	Phase # (1-16)	Set when either solid or flashing green indication has begun. Do not set repeatedly during flashing operation.
2	Phase Check	Phase # (1-16)	Set when a conflicting call is registered against the active phase. (marks beginning of max timing)
3	Phase Min Complete	Phase # (1-16)	Set when phase min time expires.
4	Phase Gap Out	Phase # (1-16)	Set when phase gaps out but may not necessarily occur upon phase termination. Event may be set multiple times within a single green under simultaneous gap out.
5	Phase Max Out	Phase # (1-16)	Set when phase max time expires but may not necessarily occur upon phase termination due to last car passage or other features.
6	Phase Force Off	Phase # (1-16)	Set when phase force off is applied to the active green phase.
7	Phase Green Termination	Phase # (1-16)	Set when phase green indications are terminated into either yellow clearance or other permissive (FYA) movement.

For the emulation of HRC data using simulation, the traffic status, signal status, and detector status in 0.1-second resolution for the analysis period are exported from the Vissim simulation. Later all these three types of data are aggregated based on the time stamp using Python programming language. The aggregated data set is used to estimate the attributes such as green occupancy ratio (GOR), red occupancy ratio (ROR), and served volume over capacity ratio (v/c).

4.4 Data Preparation

Some of the data is directly used in the analysis, while some of the data is needed further processing to make it useable for the analysis because each method requires a specific format of the data. In the following section, the data preparation process for each method is described.

4.4.1 Data Preparation for Critical Route Prediction

All data are filtered and cleaned by removing weekends, holidays, and missing data from the dataset to prepare the data for the analysis. The travel time for each minute is aggregated to 15-minute intervals starting from the beginning of the incident. For example, if the incident begins at 08:10 am, the next 15-minute data is from 08:10 am to 08:25 am and aggregated accordingly. This data is then associated with traffic attributes and incident attributes for these intervals. The processing is done using the Python programming language, which is eventually produced a dataset with all essential attributes. The combined dataset is further formatted to the specific format necessary to feed the LSTM method. The incident attributes, I-95 traffic detector, and travel time data for the I-95 segment and alternative routes for each 15-minute interval from the beginning of the incident are used as the input to the model. The model provides predicted travel time for each 15-minute interval on the alternative routes up to 90 minutes after the incident as the output in the output layer of the LSTM model.

Following the same procedure, another separate dataset for all the routes during the normal time period (excluding the incident period, weekends, and holidays) in the analysis horizon is processed to develop an LSTM model for the normal time period.

4.4.2 Data Preparation for Congestion Pattern Identification

CV, HRC, and midblock loop detector, these three types of data are collected from the case study intersection modeled in the Vissim simulation. Initially, the data for each lane is collected and aggregated for each movement group for each cycle. After aggregating, the following attributes are estimated for use in the analysis.

- *Average Travel Time:* Travel time data are collected by averaging the travel times of CVs in each cycle.
- *CV Position in Queue:* The furthest stopping position of the CVs in the queue in each cycle is collected. If no CV is determined to be stopped in the queue, this value is considered as zero.
- *v/c (Served Volume/Capacity):* Total volume of vehicle served during the green time in each cycle divided by the capacity. The capacity is estimated for normal conditions using the saturation headway after calibrating the simulation model.
- *Green Occupancy Ratio (GOR):* The ratio of the detector occupancy during the green phase to the total green time.
- *Red Occupancy Ratio (ROR):* Ratio of the detector occupancy during the first five seconds after the end of yellow in the split.
- *Mid-Block Volume:* Number of vehicles passing the midblock in each cycle.

4.5 Summary

This study utilizes both real-world and simulated data to perform the analysis. The real-world data is retrieved from RITIS and FDOT District IV, and third-party vendors. The data are then cleaned, processed, and formatted using Python programming language.

On the other hand, the simulated data is emulated using the Vissim simulation model. A real-world traffic network is used to develop and demonstrate the methodology.

CHAPTER V

MODEL APPLICATION AND RESULTS

This chapter demonstrates the application of the methodologies developed in Chapter III utilizing the data described in Chapter IV. The chapter is arranged into three major sections - the first section covers the results of the prediction of critical routes, the second section discusses the MRM modeling results and the evaluation of traffic signal plan for critical routes, and the final section includes the results of the use of CV and HRC for traffic management.

5.1 Prediction of Critical Routes

This section provides the descriptive statistics of the data, analyzes the model accuracy, and discusses the models' results for the I-95 Corridor case study.

5.1.1 Incident Statistics

The summary statistics of the incident attributes used in the analysis are shown in Table 5-1. The utilized data includes a total of 700 incidents with lane blockages that occurred during the analysis horizon in the southbound (SB) direction of I-95 from Glades Rd to Oakland Blvd. The majority of the incidents (65%) blocked one lane, and around 24% blocked two lanes of I-95. 47% of the lane blockage durations were 15 minutes or less, but there was still a significant percentage of incidents (13%), which blocked at least a lane for more than sixty minutes. Around 76% of the incidents were with low severity, around 16% were with a medium severity, and the remaining 8% were highly severe. The "Time Period" was categorized to one to four, indicating the AM Peak, Midday, PM Peak, and Evening, respectively, to use as model input. The incidents' location was measured in miles from a reference point with the mean distance equal 9.2 miles ranging from a minimum of 2.4 miles to a maximum of 16.3 miles.

Table 5-1: Incident Summary Statistics

Time Period	Total	Number of Lane Blockage			Lane Blockage Duration (minutes)					Incident Severity		
		One	Two	Three or More	≤15	>15 - ≤30	>30 - ≤45	>45 - ≤60	>60	Severity I (Low)	Severity II (Medium)	Severity III (High)
AM Peak	108	73	23	12	62	19	11	5	11	81	22	5
Midday	320	207	82	31	146	87	39	19	30	247	62	12
PM Peak	167	118	36	13	90	49	20	3	5	146	16	5
Evening	105	56	31	18	31	20	9	4	40	56	13	35
Total Incidents	**700**	**454**	**172**	**74**	**329**	**175**	**79**	**31**	**86**	**530**	**113**	**57**

5.1.2 Descriptive Statistics of Traffic Data

The descriptive statistics of the traffic input variables during the incident and normal conditions are shown in Table 5-2. The average speed, traffic count per lane, and travel time for each segment of I-95 downstream of the off-ramps to the alternative routes were used separately as model inputs to consider the effect of incident location on the alternative routes. As shown in Table 5-2, mean speeds at the freeway and alternative route locations were lower during incident conditions for all segments of I-95 and alternative routes. Moreover, travel time variations during the incidents are higher than those during normal conditions due to the variability of the incident severity. The inputs and outputs of the LSTM model are shown in Table 5-3. As mentioned previously, incident attributes, traffic attributes, and travel time in the alternative routes are used as input to the LSTM model, while the prediction of the travel times the output of the model. Among the incident attributes, three crucial parameters: Number (No) of Lane Blockage, Severity of the Incident, and Location of the Incident, are utilized in the model to capture the scenarios of the incidents in the travel time prediction on the alternative routes. The accuracy of the data, including the location and timestamps of the incidents and lane blockages, is important. Erroneous data can affect the results. However, it was determined that the quality of the FDOT incident data in the case study is sufficient for the study purpose.

Table 5-2: Statistics of Traffic Inputs during Incident and Normal Conditions

Input Variables	Description	Incident Conditions				Normal Conditions			
		Mean	Std	Min	Max	Mean	Std	Min	Max
Average speed on I-95 (mph)	Between Glades Rd and Palmetto Park Rd	51.7	0.7	10.4	77.3	60.5	0.2	10.3	79.3
Average speed on I-95 (mph)	Between Palmetto Park Rd and Hillsboro Blvd	46.7	0.5	11.0	70.7	55.3	0.2	9.9	74.3
Average speed on I-95 (mph)	Between Hillsboro Blvd and SW10th St	58.7	0.7	10.3	82.5	69.7	0.2	21.4	83.8
Average speed on I-95 (mph)	Between SW 10th St and Sample Rd	60.9	0.5	10.7	77.5	64.7	0.2	12.9	79.1
Average traffic count per lane on I-95 (per 15 mins)	Between Glades Rd and Palmetto Park Rd	235	2	107	350	267	0.3	180	358
Average traffic count per lane on I-95(per 15 mins)	Between Palmetto Park Rd and Hillsboro Blvd	366	3	116	492	380	0.3	260	482
Average traffic count per lane on I-95(per 15 mins)	Between Hillsboro Blvd and SW10th St	336	3	131	486	355	0.3	270	500
Average traffic count per lane on I-95(per 15 mins)	Between SW 10th St and Sample Rd	365	3	129	496	398	0.3	242	518
Average travel time on I-95 (sec)	Between Glades Rd and Palmetto Park Rd	45.6	1.1	24.6	261.8	34.9	0.3	24.8	262.1
Average travel time on I-95 (sec)	Between Palmetto Park Rd and Hillsboro Blvd	122.1	2.3	70.4	545.8	93.3	0.5	67.1	792.3
Average travel time on I-95 (sec)	Between Hillsboro Blvd and SW10th St	27.6	0.6	16.2	193.8	22.8	0.2	15.3	152.5
Average travel time on I-95 (sec)	Between SW 10th St and Sample Rd	109.0	2.4	76.0	1,317.4	99.0	0.7	72.4	1,089.7
Average travel time on Glades Rd (sec)	From I-95 Off-Ramp to S Military Trail	152.5	1.6	84.5	523.5	139.2	0.5	81.4	293.2
Average travel time on Palmetto Park Rd (sec)	From I-95 Off-Ramp to S Military Trail	71.4	1.0	39.2	211.7	60.7	0.2	36.1	202.1
Average travel time on Hillsboro Blvd (sec)	From I-95 Off-Ramp to S Military Trail	119.2	1.2	57.9	294.5	109.5	0.4	46.2	243.0
Average travel time on SW 10th St (sec)	From I-95 Off-Ramp to S Military Trail	93.5	1.3	40.3	248.4	77.5	0.4	40.1	252.7

Table 5-3: Inputs and Outputs of the LSTM model

Inputs	Outputs
○ No. of Lane Blockage ○ Severity of the Incident ○ Location of the Incident ○ Time of Day ○ Average Speed on Freeway ○ Average Volume on Freeway ○ Travel Time on Freeway ○ TimeStep* ○ Travel Time on Alternative Route I ○ Travel Time on Alternative Route II ○ Travel Time on Alternative Route III ○ Travel Time on Alternative Route IV	○ Travel Time on Alternative Route I ○ Travel Time on Alternative Route II ○ Travel Time on Alternative Route III ○ Travel Time on Alternative Route IV

*TimeStep= Time after an incident; For example, 15 minutes after the occurrence of the incident was termed TimeStep 1, and 30 minutes after the occurrence of the incident was termed as TimeStep 2, and so on.

5.1.3 Evaluation of the Travel Time Prediction Model

Two performance metrics were used to evaluate the performance of the trained LSTM models: Mean Absolute Deviation (MAD) and Mean Absolute Percentage Error (MAPE). These measures were estimated using the test dataset for the case study described earlier. The performance matrices for each timestep for the two LSTM models (the incident period model and the normal period model) for each investigated diversion segment are shown in Tables 5-4 and 5-5, respectively. It is observed from the tables that the accuracy of the prediction is good across the timesteps. However, the model for the normal period has lower errors compared to the model for the incident period because of the availability of a higher number of normal period data points to train the model and lower variability in the travel time since the congestion conditions due to incidents vary depending on incident attributes. The MAD of travel time for the four routes ranges from 12 seconds to 20 seconds during the incident and from 6 seconds to 14 seconds during the normal period. Overall, the full model during the incident period is able to predict the travel time with accuracy between 82% and 90%, while during the normal period, the accuracy range is between 88% and 92%. Considering the high stochasticity of the travel time and other traffic data, the prediction accuracy of the models is acceptable.

Table 5-4: Model Performance Statistics for the Incident Period Prediction

Error Measure	Alternative Segments	Timesteps						Full Model
		1st Step	2nd Step	3rd Step	4th Step	5th Step	6th Step	
MAD (sec)	Glades Rd	12.8	14.1	13.5	13.5	14.3	15.0	20.8
	W Palmetto Park Rd	18.3	16.9	17.1	17.8	16.8	20.2	13.5
	W Hillsboro Blvd	15.9	15.1	15.9	15.7	15.9	13.7	19.8
	SW 10th St	13.9	16.2	14.6	13.9	16.0	17.0	15.0
MAPE (%)	Glades Rd	21.1	26.3	22.0	20.3	26.2	26.6	13.2
	W Palmetto Park Rd	14.6	12.9	13.9	12.6	13.7	16.6	17.8
	W Hillsboro Rd	19.3	20.9	19.7	19.4	20.6	19.3	15.7
	SW 10th St	13.4	15.5	13.7	12.4	17.1	16.4	15.4

Table 5-5: Model Performance Statistics for the Normal Period Prediction

Error measure	Alternative Segments	Time Steps						Full Model
		1st Step	2nd Step	3rd Step	4th Step	5th Step	6th Step	
MAD (sec)	Glades Rd	14.2	10.6	10.5	12.1	10.9	11.1	12.4
	W Palmetto Park Rd	9.1	6.5	6.1	6.3	6.1	6.0	6.9
	W Hillsboro Blvd	14.9	11.8	12.1	11.7	11.5	11.5	13.4
	SW 10th St	9.6	7.6	7.6	9.4	8.1	8.3	9.5
MAPE (%)	Glades Rd	9.7	7.8	7.3	7.6	7.5	7.5	8.6
	W Palmetto Park Rd	14.3	10.3	9.9	9.8	9.7	9.5	10.8
	W Hillsboro Rd	13.5	11.1	11.1	10.5	10.3	10.4	12.5
	SW 10th St	12.2	10.7	10.7	11.3	10.9	10.9	12.1

5.1.4 Travel Time Change on Alternative Routes

The percentage change in travel time on the potential alternative routes (ΔTT) for all timesteps for all the incidents was estimated based on the results from running the two models for each incident. The cumulative distributions of ΔTT for all routes for the first three timesteps are shown in Figures 5-1 to 5-3. Due to the incidents, the travel times on the alternative routes increased by up to 150% after 15 minutes (Figure 5-1), 250% after

30 minutes (Figure 5-2), and 200% after 45 minutes (Figure 5-3) of the incident occurrence. One can see that 50% to 80% of the predicted travel times using the LSTM model with incident conditions have travel times within 25% of the values predicted by the LSTM model for no-incident conditions. The variations in travel times during these incidents within 25% of the "normal" travel time reflect the normal day-to-day variations in travel time, and thus no diversion is assumed for these incidents. For the demonstration, the increase in travel time by more than 25% on the alternative routes is assumed to require the agency's attention for the case study. However, as stated earlier, the threshold value for ΔTT can be set based on the needs and resources of the local agencies to identify the critical routes that require attention.

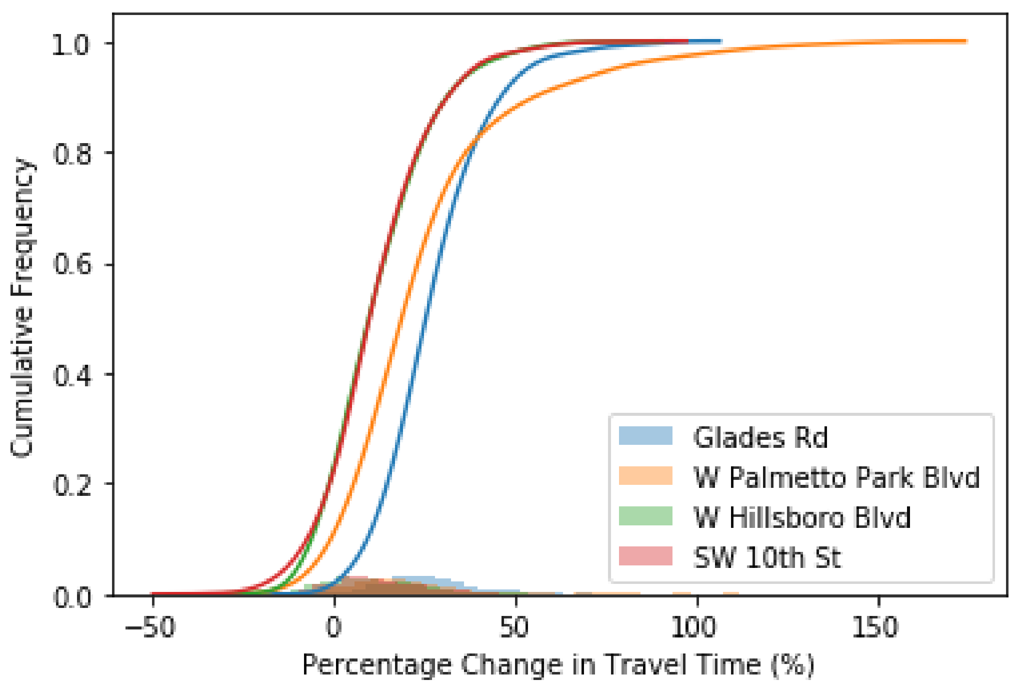


Figure 5-1: Percentage Change in Travel Time after 15 Minutes of the Incident using the Developed LSTM Models

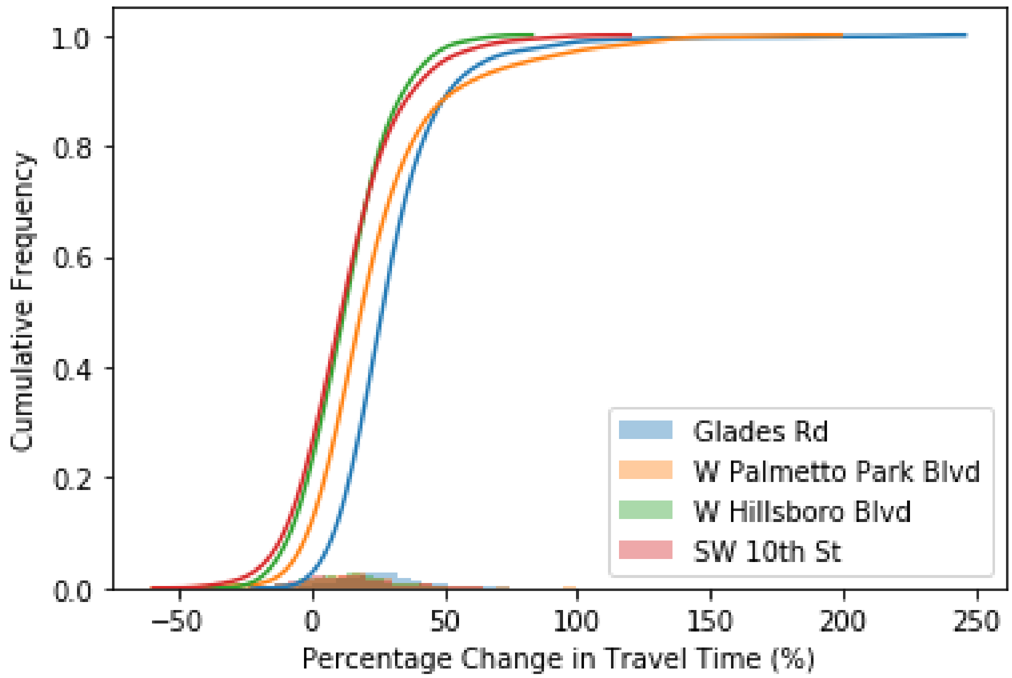


Figure 5-2: Percentage Change in Travel Time after 30 Minutes of the Incident using the Developed LSTM Models

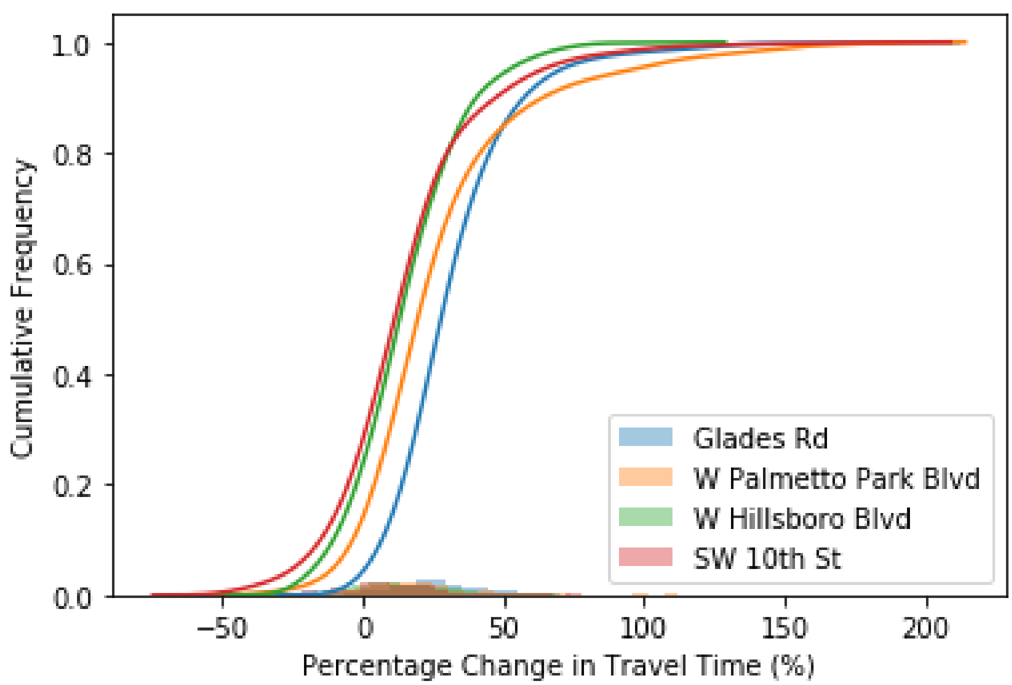


Figure 5-3: Percentage Change in Travel Time after 45 Minutes of the Incident using the Developed LSTM Models

5.2 MRM and Traffic Signal Plan Development and Evaluation

This section presents the clustering results and analysis for the selection of representative scenarios for MRM modeling, analysis of MRM modeling, and development and evaluation of traffic signal plans.

5.2.1 Clustering Analysis

Clustering analysis results are divided into two sections: i) determination of the optimum number of clusters and ii) selection of representative scenarios from the clusters. Determination of the optimum number of clusters is particularly crucial for the analysis as it identifies the unique patterns present in the dataset.

5.2.1.1 Determination of Optimum Number of Clusters

The optimal number of clusters was determined using the Elbow method. Initially, the t-SNE method was used to visualize the data to see the intrinsic patterns in the dataset. As stated earlier, t-SNE reduces the dimension of the data into two dimensions. Figure 5-4 shows the reduction of the ΔTT on the W Palmetto Park Rd for all 15-minute timesteps after the occurrence of all incidents to two dimensions using the t-SNE method. The figure depicts the presence of intrinsic clustering patterns in the dataset. In the Elbow method, the sum of square error (SSE) was plotted against the number of clusters obtained using the K-means clustering, as shown in Figure 5-5. The location of the kink in the elbow in Figure 5-5 indicates that twelve clusters are the optimal number of clusters to represent the impacts of diversion on the alternative routes during incidents. Twelve clusters were further analyzed next to identify the distinct scenarios for the signal control plan development.

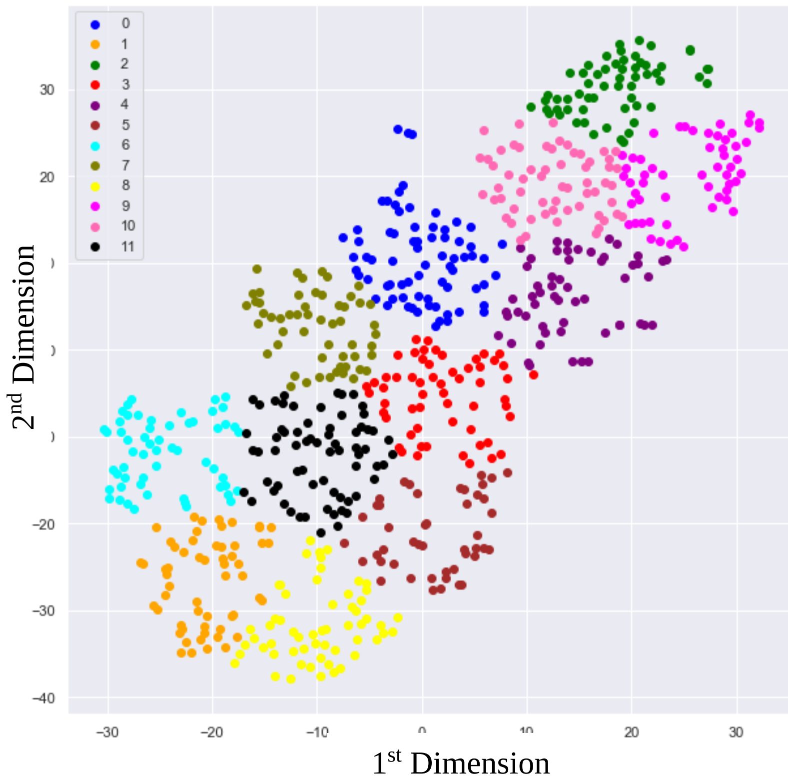


Figure 5-4 Visualization of Data in Low Dimensions using t-SNE

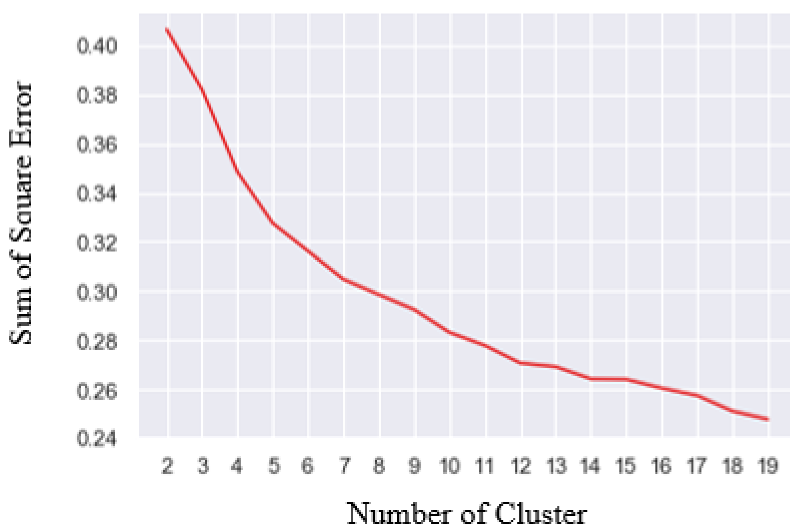


Figure 5-5: Elbow Plot Selection of Scenarios for Plan Development

5.2.1.2 Clustering Results and Selection of Representative Scenarios

Table 5-6 shows the average *ΔTT* for the 12 selected clusters for the six 15-minute timesteps after the incidents. The analysis of the clusters indicates that the average *ΔTT* in Cluster 1 and Cluster 8 for all timesteps is within 25% of the normal day's travel time and is within the natural variation of the day-to-day traffic. Thus, no new signal plans were developed for these two clusters, and they were excluded from further analysis. The remaining ten clusters show distinct patterns of *ΔTT*. Patterns on these clusters do not only vary across the clusters but also across the timesteps. This is expected to happen because some incidents cause the immediate diversion of traffic to the examined alternative route, while other incidents induced diversion at a later stage of the incidents based on the incident and traffic characteristics such as incident location relative to the freeway off-ramp exit to the alternative route, severity, number of lane blockage, traffic demands, and so on. The incident characteristics associated with each cluster are also shown in Table 5-6. For Clusters 1, 8, and 10, the incidents mostly occurred during the Midday or Evening periods, the severity of the incidents was low, and the locations were far from the exit to the alternative routes. The effect of the incidents on the alternative route travel times was very high when it happened during the AM or PM peaks and close to the off-ramp exit to alternative routes, as in Clusters 2 and 3. Medium to high severity incidents grouped in Clusters 5, 11, and 12 affected the alternative routes during AM, PM peak, and Midday when the location of the incidents was far from the alternative routes. Cluster 6 reflects low to medium severity incidents during the PM peak, and Cluster 7 includes diversion during the Midday when the incidents occurred close to the exit to the diversion routes.

The variation of the average ΔTT across timesteps and clusters, as shown in Table 5-6, requires different signal timing plans to accommodate the varying diverted traffic demands and impacts. Typical incidents associated with each of the clusters were coded in the Visum mesoscopic model to simulate the associated scenarios and utilize the dynamic traffic assignment of the model to estimate the demands on each link of the alternative routes. The demands were then imported from Visum to Vissim to allow the optimization of the signal timing plan for each scenario within the microscopic model environment, as described earlier. Please, note that although the clustering results are presented for all time periods in Table 5-6, the modeling and the optimization analyses presented in the remaining of the study are only for the AM peak period.

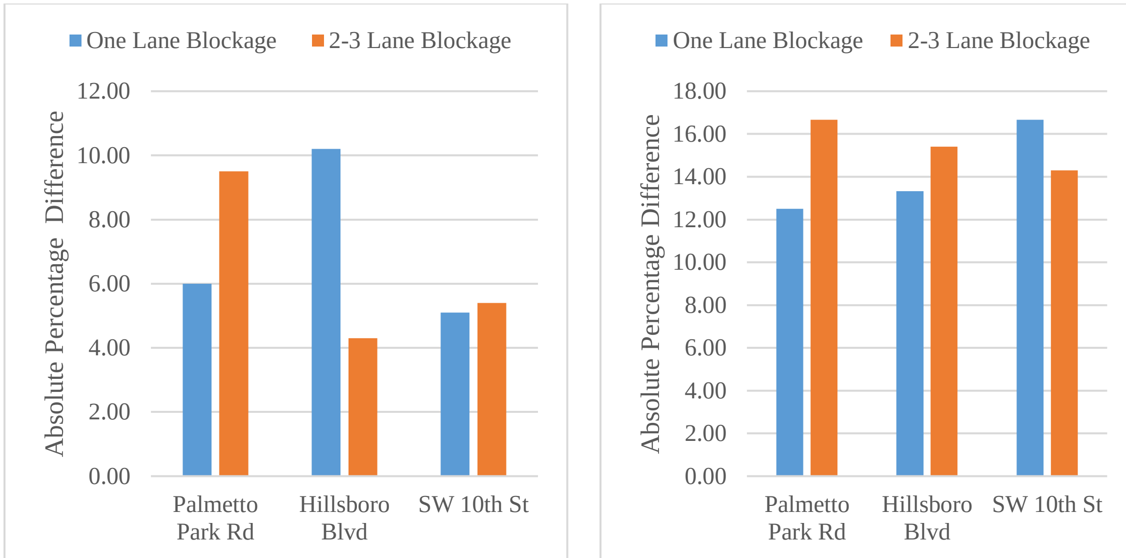
Table 5-6: Average Percentage Change in Travel Time and Signal Plan for Scenarios

	Avg. ΔTT (%) in Different Timesteps						Incident Characteristics	Analysis Scenarios
	1st Step	2nd Step	3rd Step	4th Step	5th Step	6th Step		
Cluster 1	X	X	X	X	X	X	Period: Midday & Evening Location: Far from the alternative routes Severity: Low	
Cluster 2	33.7	39.7	79.9	193.1	210.2	211.2	Period: AM and PM peak Location: Close to the alternative routes Severity: Low	Scenario I
Cluster 3	103.4	60.4	49.4	32.6	32.5	X	Period: AM and PM peak Location: Close to the alternative routes Severity: Medium to high	Scenario II
Cluster 4	29.2	62.3	82.1	27.8	X	X	Period: AM and PM peak Location: Far from the alternative routes Severity: Low	Scenario III
Cluster 5	X	X	28.2	43.6	31.3	147.5	Period: Midday Location: Close to the alternative routes Severity: Medium to high	
Cluster 6	103.2	126.3	136.2	126.7	67.4	65.8	Period: PM peak Location: Far from the alternative routes Severity: Low to medium	
Cluster 7	X	29.5	36.8	59.5	125.5	111.1	Period: Midday Location: Close to the alternative routes Severity: Low	
Cluster 8	X	X	X	X	X	X	Period: Midday & Evening Location: Far from the alternative routes Severity: Low	
Cluster 9	60.8	X	X	X	X	X	Period: AM and PM peak Location: Far from the alternative routes Severity: Low	Scenario IV
Cluster 10	X	X	25.1	28.2	27.5	30.1	Period: Midday & Evening Location: Far from the alternative routes Severity: Low	
Cluster 11	49.0	70.1	100.0	133.0	125.2	71.3	Period: AM and PM peak Location: Far from the alternative routes Severity: Medium to high	Scenario V
Cluster 12	51.4	46.1	47.7	46.4	54.9	61.5	Period: Midday Location: Far from the alternative routes Severity: Medium to high	

Note: ‘X’ ≤ 25%

5.2.2 MRM Model Calibration Results

A significant aspect of this study is to calibrate the simulation models not only to reflect normal conditions but also the diversion during incident conditions. The calibration results of the mesoscopic model for the normal condition meet the FHWA Traffic Analysis Toolbox (TAT) Volume III and are not presented here. More interesting are the results of calibration of the model for the incident conditions considering the travel times and path-based demands on the diversion routes, as obtained based on data from third-party vendors. These variables' values for the three-diversion links between I-95 and S Military Trail were close to the FHWA TAT Volume III criteria. Figure 5-6(a) shows that the difference between the model and real-world travel times for one-lane blockage and 2-3 lane blockage incidents were below 15%, as specified by the FHWA TAT Volume III. In the case of path-based traffic (Figure 5-6(b)), the modeled volume of the SW 10th St link for one-lane blockage incidents and the Palmetto Park Rd link for 2-3 lane blockage incidents were slightly over 15%.



a) Difference in Travel Time **b) Difference in Path-Based Traffic**

Figure 5-6: Modeled and Real-World Travel Time and Path-Based Traffic Difference during Incidents

5.2.3 Traffic Signal Plan Development and Evaluation

This section describes the results of traffic signal plan generation and evaluation of the plans.

5.2.3.1 Pareto Front

The Pareto front, which is used in the optimization of the signal timing in this study, considers a set of non-dominated solutions to achieve an optimal trade-off between the competing objectives. The Pareto fronts in the signal timing optimization of all scenarios are shown in Figure 5-7. The fronts in this study consist of two competing objectives: average delay and overall throughput. The Pareto fronts for different plans moved upward compared to the Pareto front for the normal conditions, as the developed solutions for the incident diversion scenarios were able to increase the throughput without adversely affecting the average delay. The solutions at the two ends of each front signify the two extreme solutions corresponding to their objectives. Although the solutions in the middle of the front are optimal solutions based on both objectives, in special scenarios, agencies may decide to prioritize one objective over another, such as prioritizing the throughput on the alternative route compared to the total delay of the intersections.

The movement of the Pareto front from the initial generation to the final generation of the Genetic Optimization of the signal timing plan associated with Scenario IV is shown in Figure 5-8. The approximated Pareto front for five selected generations shows the improvement of the solutions from one generation to the next. Likewise, the solutions in the final generation are significantly better than the solutions of the first generation in terms of the objectives functions' values and component variables, which confirms the success of the multi-objective optimization.

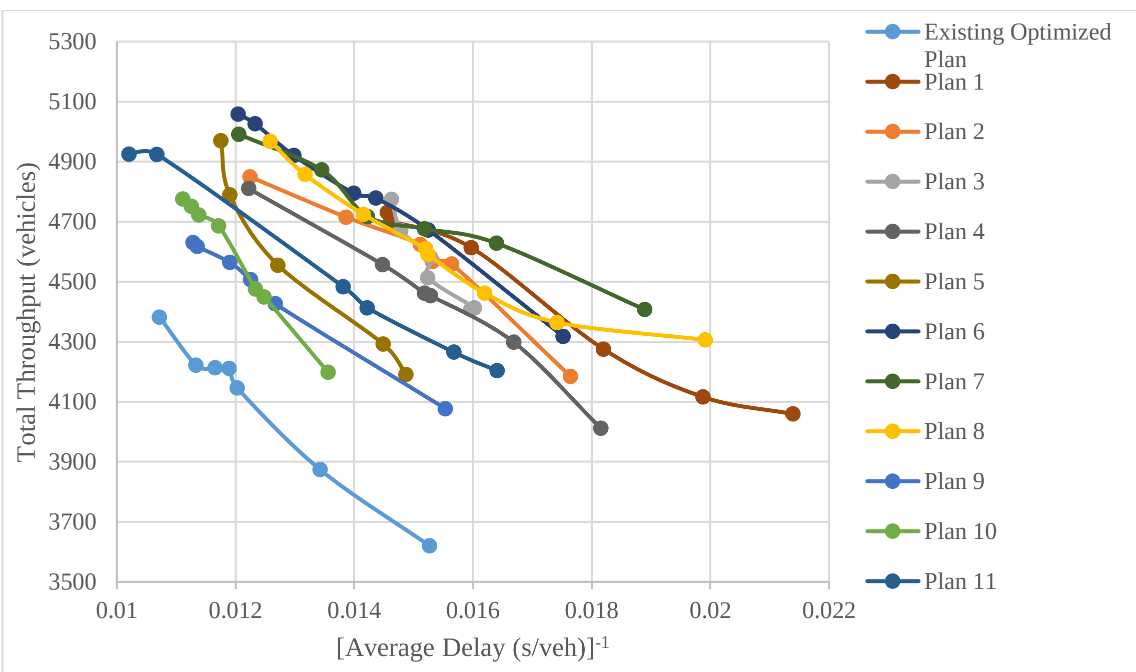


Figure 5-7: Approximated Pareto Front for All Scenarios

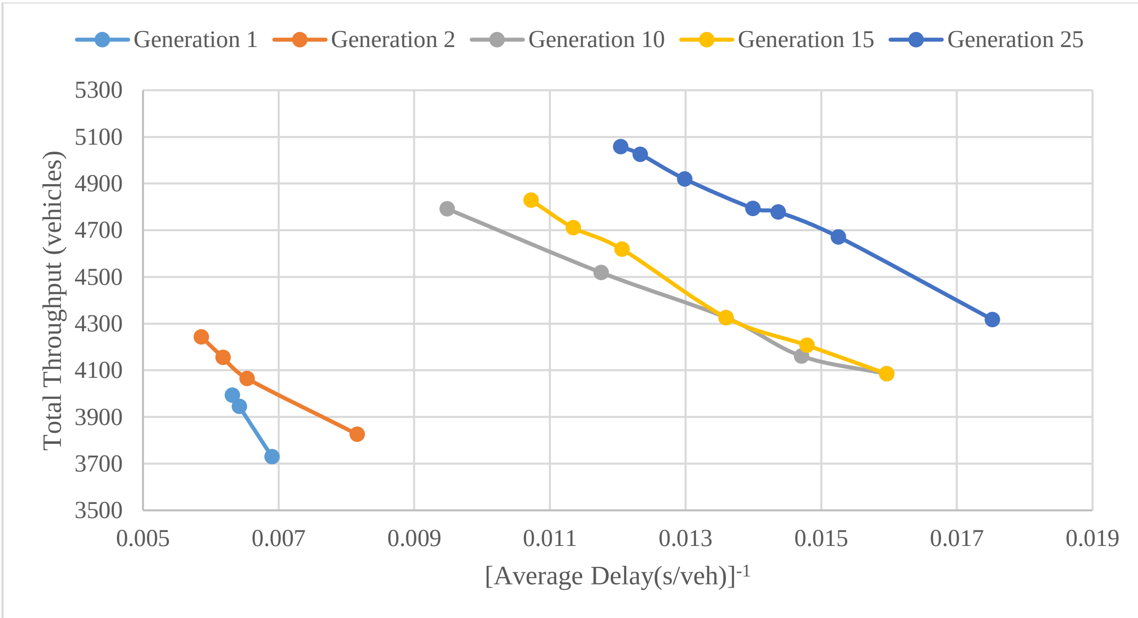


Figure 5-8: Pareto Front for Different Generations

5.2.3.2 Evaluation of the Plans

The optimized plans for each scenario were evaluated to assess their performance using the microscopic simulation model. For this comparison, the plan that gives the maximum throughput for each scenario was selected from their corresponding Pareto front. This was done to accommodate the demand surge resulting from the diversion. The plan that provides the maximum throughput during the normal condition (without considering the diverted demand) was identified from the corresponding Pareto front as the base scenario for comparison. Evaluation of the plans was performed at both the network and diversion route movement levels. The diversion route movements included in the comparison are the west-bound left-turn movement at the W Palmetto Park Rd-S Military Trail intersection and the south-bound through movements of two downstream intersections: Hillsboro Blvd-S Military Trail and SW 10th St-S Military Trail.

Figure 5-9 shows the percentage changes in both delays and throughputs for the newly developed plans for all scenarios. The developed plans for all scenarios increased the throughput while reducing the overall delay compare to the values obtained for the base scenario plan. However, the performance improvement for the diversion route movements was far more significant than those for the overall network. For Scenario I, the throughput increased by 13% and 72% with the optimized plans compared to the base scenario plan for the whole network and the diversion route movements, respectively. For the same scenario, the overall delay for the entire network and the diversion route movements was reduced by 17% and 54%, respectively. The delay and throughput changes for Scenario IV were the lowest as the diversion impacted one timestep only. The increase in throughput and reduction in delay were higher for Scenario II than those for Scenario III due to the longer duration of impact on the diversion route in Scenario II. Although all six timesteps were impacted due to diversion in Scenarios I and V, the reduction in delay and increase in throughput were higher for Scenario I than those for Scenario V because of the high severity of the impact in the case of Scenario I.

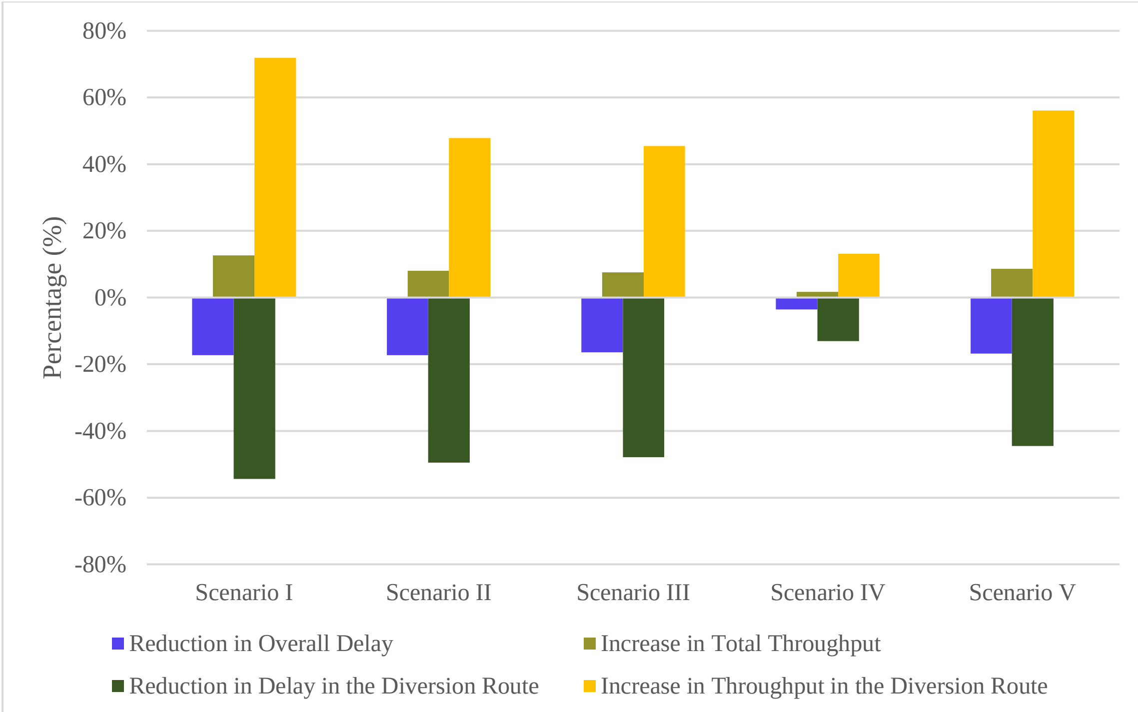
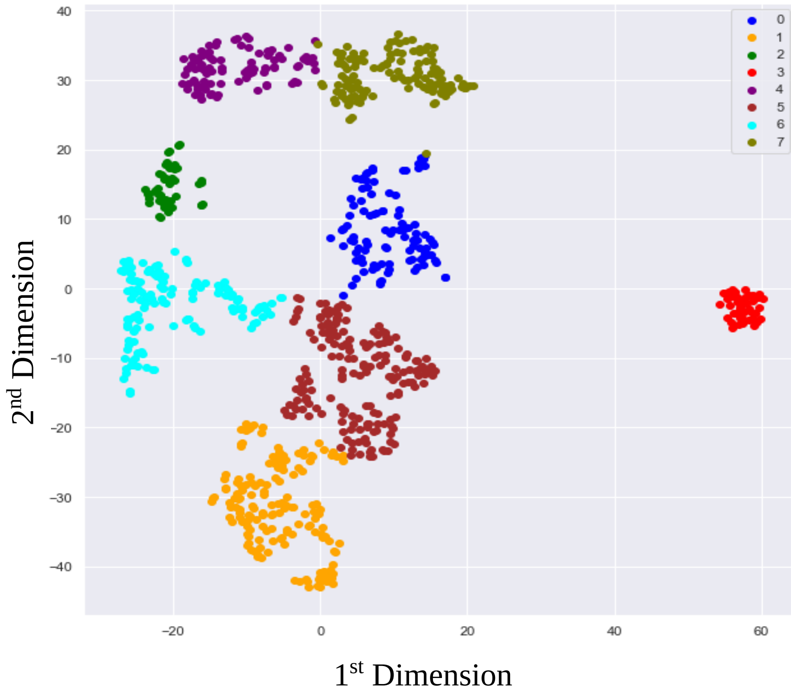


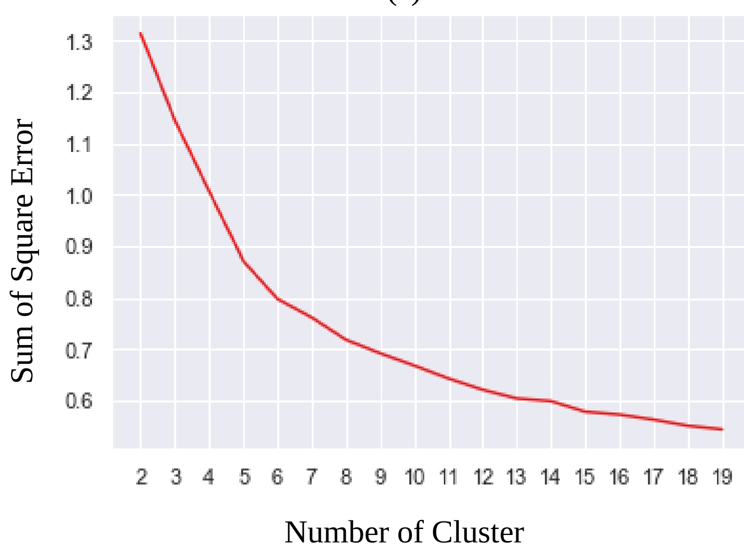
Figure 5-9: Evaluation of Derived Traffic Signal Control Plans for Scenarios

5.3 Congestion Identification using CV and HRC

Congestion at the intersection level was identified using clustering analysis. Clustering methods were applied to the data attributes obtained using CV and HRC data. The optimum number of clusters was determined by visualizing the dataset in low dimensions using t-SNE, which was further verified using the Elbow method (Ketchen and Shook, 1996). Initially, the data were reduced to two dimensions using t-SNE and visualized (Figure 5-10(a)). The figure depicts eight different patterns, color-coded in different colors. Although the interpretation of the cluster is not feasible from the figure, the method was found successful in demonstrating the number of clusters present in the dataset. Then, utilizing the Elbow method, the sum of square error (SSE) was plotted against the number of clusters using the K-means clustering, as shown in Figure 5-10(b). Based on the location of the kink in the elbow, the figure also recommends eight clusters as the optimum number of clusters. Therefore, eight clusters were considered as the optimum number and used across the methods to evaluate the performance of the methods in identifying congestions.



(a)

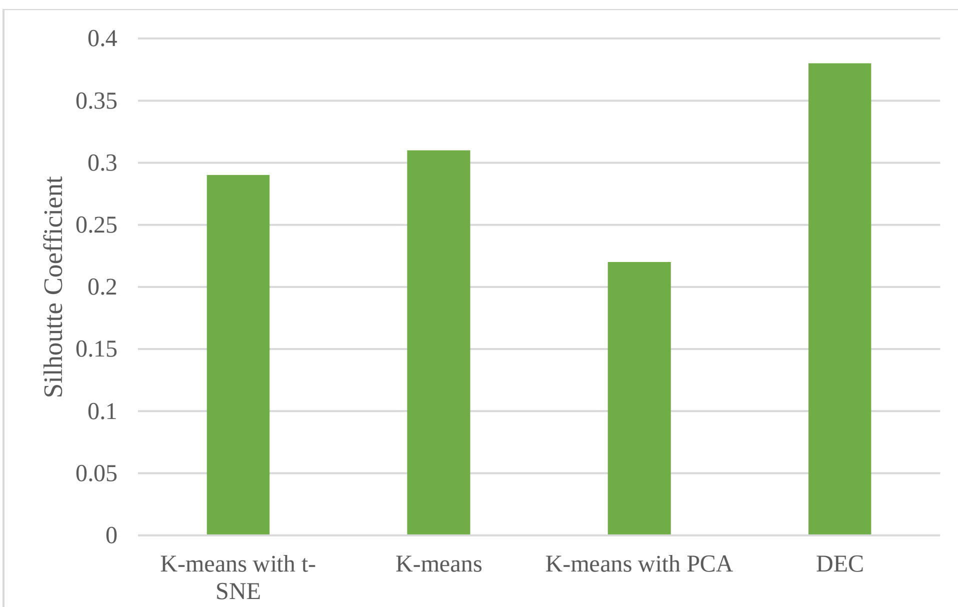


(b)

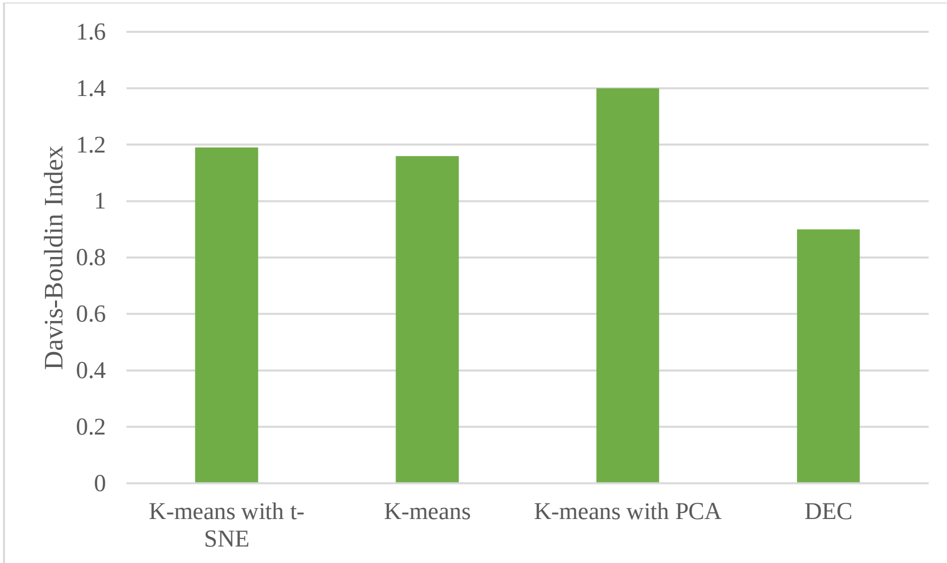
Figure 5-10: (a) Visualization of the Data in Low Dimensions using t-SNE, (b) Sum of Square Error for Different Number of Clusters

5.3.1 Performance Evaluation of Clustering Methods

As stated earlier, four different clustering techniques (two linear and two non-linear) were applied, and their performances were evaluated. The values of the Silhouette Coefficient and the Davies–Bouldin index (DBI) from applying the methods are shown in Figure 5-11(a) and Figure 5-11(b), respectively. Based on the measures (the Silhouette Coefficient and the DBI), the DEC method came out as the best method among the four. DEC produced dense, appropriate, and well-separated clusters compared to the other methods. The performance of the K-means with PCA was the worst among the methods because of the non-linear relationship among the attributes of the data. However, the K-means method clustered the data better than the K-means with t-SNE. This possibly happened because the t-SNE approach does not preserve the distances or density like its linear counterpart (the PCA) during the dimension reduction.



(a)



(b)

Figure 5-11: (a) Silhouette Coefficient, (b) Davis-Bouldin Index (DBI) for Different Clustering Methods

5.3.2 Congestion Patterns Identification

The eight clusters identified using the DEC method for 100% CV market penetration were evaluated to identify the congestion patterns and associated characteristics of the congestions. The features of the clusters are shown in Table 5-6. Average queue length, average travel time, and green occupancy ratio (GOR) for both thru and left-turn movements in each cluster are shown in Table 5-7. Clusters 2 and 3 represent normal conditions of the intersection among the clusters. However, the average queue lengths and travel times are higher in Cluster 3 than Cluster 2. Cluster 0 represents an extreme congested condition for both movements as indicated by the high travel time and queue length; however, the low GOR of the left-turn movement represents starvation of vehicles of that movement. This has happened because the oversaturated thru movement vehicles blocked the entrance to the left turn bay and thus prevented the left-turning vehicles from entering the left-turn bay. Therefore, Cluster 0 represents extreme congestion of the thru movement. Clusters 5 and 6 represent the opposite of Cluster 0, where the left-turn movement is highly congested. The oversaturated left-turn vehicles filled the left-turn bay and spilled over to the adjacent thru movement lanes, thus restricting the thru vehicle's movements. Clusters 1 and 7 represent slightly and moderately affected left-turn movement, respectively. In these cases, the left turn vehicles did not affect the thru movement. Cluster 4 represents a moderately impacted left turn that affects the thru movement. From Table 5-7, it is also observed that the queue length of the left turn up to 500 ft does not affect the thru movement. However, the increase in the left turn queues beyond that number affects both movements severely. In this situation, providing more green to the left turn movement will be able to reduce the overall delay at the intersection.

Table 5-7: Characteristic of the Clusters and Associated Patterns of Congestion

Clusters	Thru Movement			Left Turn Movement			Remarks
	Avg. Queue Length (ft)	Avg. Travel Time (sec)	GOR	Avg. Queue Length (ft)	Avg. Travel Time (sec)	GOR	
Cluster 0	2,432.1	207.6	0.88	2,425.6	300.3	0.5	Extremely affected thru movement
Cluster 1	216.6	66.3	0.65	424.1	197.6	0.89	Slightly affected left movement
Cluster 2	201.1	68.7	0.57	168.2	77.7	0.62	Normal condition
Cluster 3	235.3	70.8	0.67	202.7	79.3	0.61	Moderately normal condition
Cluster 4	402.4	94.6	0.68	1,954.9	716.1	0.96	Moderately highly affected left turn movement
Cluster 5	2,147.7	185.2	0.73	2,310.9	753.4	0.98	Highly affected left turn movement
Cluster 6	905.2	127.2	0.51	2,417.5	1,090.2	0.98	Extremely affected left turn movement
Cluster 7	210.4	64.9	0.57	507.8	217.1	0.89	Moderately affected Left turn

5.3.3 Evaluation of the Impact of CV Market Penetration (MP)

The impact of CV MP in determining the accurate clusters vis-à-vis congestion patterns with respect to base condition (100% CV MP) was evaluated, and the results are presented in Figure 5-12. The figure demonstrates high accuracy in identifying congestions with as low as 10% CV market penetration. The accuracy with 10% CV is around 90% and increased to around 98% when the CV MP is 90%. The accuracy is about 95% when the CV percentage increased to 20% and remained almost the same, up to 40% MP. To check whether the methodology is applicable in the case of no CV vehicle present in the traffic stream, the clustering analysis was done considering all the vehicles as conventional vehicles, and the agency does not have travel time and queue length information from the connected vehicles. Based on the accuracy measure, it is observed that without CV, the method can identify congestion with 70% accuracy. This accuracy may improve if the movement travel time data based on other means such as probe vehicles, Bluetooth technology, or detector base technology are used in the analysis.

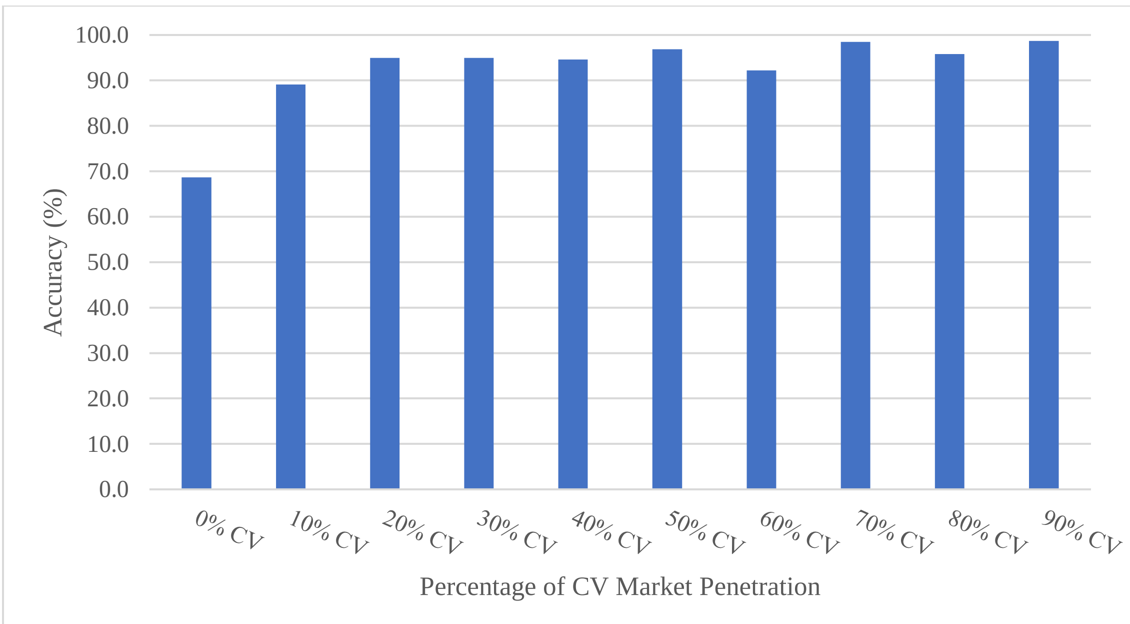


Figure 5-12: Accuracy in Predicting Congestion Patterns in Different MP of CV

5.4 Summary

In summary, applying the LSTM model to predict travel time on alternative routes using incidents, traffic, and travel time attributes shows a great promise and generates results with acceptable accuracy. Thereby, the method is suitable to predict critical alternative routes dynamically after the incident in the freeway. The *ΔTT* value used as threshold measures for identifying the alternative routes shows that percentage changes in the travel time remain within 25% for 50 to 80% of the incidents. This is caused by the normal day-to-day variation of the traffic and can be assumed very low diversion and does not require further attention. For other cases, the traffic management agency can implement special signal control plans.

An MRM modeling framework for estimation of path-level traffic and traffic signal plans for the diverted routes are developed. Calibration of the mesoscopic model for incident conditions produces path-level traffic in the diversion routes. Special traffic signal plans developed using microscopic simulation for the corresponding path-level traffic are assessed against the existing TOD plans for various scenarios and found effective in reducing delay and increase in throughput both network-wide and in the diversion routes.

Congestion identification using HRC and CV data show that very low market penetration of CV with HRC could estimate the congestion at the intersection level with greater accuracy.

CHAPTER VI

CONCLUSION AND RECOMMENDATIONS

Traffic diversion during incidents is a proven technique for mitigating the congestion on the affected routes, improving the mobility of the routes and the entire corridor. A CFA operation requires proper management of diversion routes through proactive strategies since reactive strategies may not be able to improve the already deteriorated traffic condition. Proactive strategies of CFA include dynamic prediction of alternative routes and implementation of the special signal timing plan corresponding to the diverted traffic. Prediction of alternative routes and the implementation of special signal plans are particularly important for diversion management because it is a dynamic phenomenon greatly influenced by the prevailing traffic and incident attributes. The existing data sources provide the agency an upper hand to design proactive methods for CFA; however, the availability of new data sources such as CV and HRC data could improve the model results and provide better control in the operation of CFA.

6.1 Summary and Conclusion

This research aims to develop a proactive method for CFA during incidents in the freeway using existing data sources and also explores the use of CV and HRC data for this purpose. In this regard, the research is concerned with the following:

- Develop a method for predicting the critical routes utilized by motorists during incidents and the associated traffic and incident conditions under which the diversion occurs due to the incident on the freeway.

- Develop a method for generating special signal timing plans and estimate the benefits of activating those plans to mitigate the deterioration in the performance of the movements of critical route intersections due to diversion.
- Estimate the benefits of utilizing new data sources (i.e., CV, HRC) as part of the methodology to support traffic management in the network.

A major effort to achieving these objectives is involved in data collection and preparation, including network coding and simulation modeling. Traffic, travel time, and incident data retrieved from three different sources are aggregated and formatted to make them appropriate as inputs to the analyses. Models at the macroscopic, mesoscopic, and microscopic resolutions are coded using three different software (e.g., Cube, Visum, and Vissim) in an MRM modeling structure. Before utilization, all these simulation models are calibrated using the real-world data and validated using the criteria set in FHWA TAT Volume III.

With regard to predicting the utilized critical routes, the study demonstrated the effectiveness of a method to proactively identify the alternative routes utilized by traffic diverted due to incidents in real-time based on the prediction and prediction of the travel times on alternative routes. A machine learning method, LSTM, is used to predict the travel times of alternative routes based on traffic, incident, and travel time data. The travel times are predicted for both incidents and normal conditions from the detection of the incident to 90 minutes later at 15-minute intervals. The percentage of increase in travel time due to incidents was used to identify the critical alternative routes.

Results of the LSTM models show their capability in predicting travel time on alternative routes with acceptable accuracy. The developed models are dynamic and able to capture traffic and incident variation to determine the impacts on the alternative routes in real-time. The increases in travel time of the alternative routes indicated that 50-80% of the incidents do not affect the routes. The remaining incidents caused an increase in travel times, requiring a special signal plan to accommodate the diversion. The developed method provides the agency with an easily implementable operation strategy to operate both freeway and arterials coordinately during incidents and facilitate the diversion. The methodology mainly requires travel time data that are becoming available to the agency through third-party vendors or automatic vehicle identification readers like Bluetooth readers. The selection of the potential alternative arterials and impacted movements on the arterials require expert knowledge or real-time observation of navigation apps after the incidents.

For developing the special signal plans considering diversion, which is the second objective, the study develops and uses an MRM model for this purpose. The methodology identifies the impacts of diversion utilizing clustering analysis based on the increase in travel times on the alternative routes following the occurrence of incidents. The scenarios identified based on clustering are modeled utilizing an MRM modeling approach to estimate the demands on the diversion routes. The MRM is calibrated based on path-level demand data and travel time data obtained from third-party vendors. These demands are then used as inputs to microscopic-based optimization of signal timings, deriving special signal timing plans to activate in the event of diversion. The proposed multi-objective optimization method allows the agency to prioritize different objectives in the optimization based on the prevailing condition, available resources, and purpose.

The evaluation of the signal timing plans resulting from the multi-objective signal timing optimization indicates that the derived special signal timing plans are able to reduce the delays and increase the throughputs in the network, particularly for the traffic movements utilized by the diverted traffic. The degrees of improvement depend on the impact of the diverted traffic on alternative routes' operations.

The MRM approach provides the agency the opportunity of modeling different incidents, associated diversion of traffic, and the resulting impacts. The utilized approach emphasizes the importance of calibrating the percentage diversion of traffic to alternative routes and the impacts on the alternative route travel times in the mesoscopic simulation-based DTA. This study successfully demonstrated this calibration, which has not been done in the past. The emerging data sources, including those from third-party vendors, high-resolution controller data, and connected vehicles, provide the needed information for such calibration. The methodology developed in the first and second objectives can be used to support the selection of management plans as part of real-time decision support systems (DSS) at traffic management centers.

The final objective of this research is to explore the use of CV and HRC data for intersection-level congestion identification. The study demonstrates a methodology to identify the congestions at the intersection microscopically using a very low market penetration of CV combined with HRC data. Two linear clustering methods, the K-means, and K-means with PCA, and two non-linear clustering methods, the K-means with t-SNE and DEC, are investigated to identify the congestions patterns at the intersections. The DEC clustering method is found to be the best among the four clustering approaches in separating the different congestion types into different clusters. Clustering based on the

HRC data combined with as low as 10% market penetration of CV data produces a very good performance compared to ground-truth clustering (clustering with 100% market penetration of CV).

6.2 Research Contributions

During the incident, diversion of traffic is an integral part of the TIM and ICM that advocates the coordinated operation of traffic between adjacent facilities. This research developed a machine learning model to predict the critical alternative routes during the incident in the freeway that motorists will use for diversion in the next one and a half hours. The model is unique in three ways: i) this is the first model that addresses the dynamic relationship between diversion of traffic and incidents based on detailed data on the alternative routes, ii) it provides a proactive approach of traffic management through the prediction of the critical alternative routes in the analysis horizon, iii) the model is developed utilizing existing data sources available to the agency which makes it readily implementable within the existing framework of the traffic management center with minimal cost.

The study also developed a new approach to traffic signal plan development in the alternative routes that facilitate the diversion without deteriorating the intersection's overall performance. The MRM modeling, especially the mesoscopic simulation model for identifying the traffic in the alternative routes during incidents, is relatively new. The MRM approach used in this study provides the agency the opportunity of modeling different incidents, associated diversion of traffic, and the resulting impacts. The multi-objective optimization for signal control plan development provides the agency the flexibility to prioritize different objectives in the optimization based on the prevailing condition, available resources, and purpose. Overall, the entire methodology for coordinated operation of freeway and arterial during incidents eliminates the need for online simulation advocated by existing studies that many agencies are not capable of or do not have the resources to do that.

The use of HRC and CV data for congestion identification due to diversion provides a new perspective to the agency of using these imminent technologies. Identification of congestion on the intersection level is crucial for designing signal splits for each movement. The methodology developed in the study paved the way for congestion identification in low market penetration of CV with HRC. Considering the benefits of these new technologies, the agency can plan its investment policy in the HRC and CV technology.

6.3 Recommendations for Future Research

This dissertation work can be extended in the future in the following ways. The study utilized the machine learning technique for the identification of the alternative routes and used the MRM technique to estimate the path level traffic for signal control development. Future research could use the machine learning technique to estimate the path-level traffic data and compare the results with the MRM technique.

The study did not consider the effect of incidents that may occur on the alternative routes in predicting the diversion scenarios and developing the signal plans corresponding to these scenarios. Future research could explore the scenarios of simultaneous incidents on both the freeway and arterials to assess their overall impact on the corridor.

The study utilized CV data for determining intersection-level congestion due to diversion. On top of that, future studies could explore the use of CV data for the signal plan development corresponding to the congestion patterns.

REFERENCES

1. Abu-Lebdeh, G., & Benekohal, R. F. (2000). Genetic algorithms for traffic signal control and queue management of oversaturated two-way arterials. *Transportation Research Record, 1727*(1), 61-67.

2. Adler, D. (1993). Genetic algorithms and simulated annealing: A marriage proposal. *IEEE International Conference on Neural Networks* (pp. 1104-1109). IEEE.

3. Ahmed, F., & Hawas, Y. E. (2015). An integrated real-time traffic signal system for transit signal priority, incident detection and congestion management. *Transportation Research Part C, 60*, 52-76.

4. Al-Deek, H., Venkata, S., Flick, J., & Khattak, A. (2009). Dynamic message sign deployment and diversion behavior of travelers on Central Florida toll roads. *Transportation Research Record: Journal of the Transportation Research Board, 2129*, 24-34.

5. Alexiadis, V., & Armstrong, A. (2012). *Integrated Corridor Management Modelling Results Report: Dallas, Minneapolis, and San Diego.* FHWA-JPO-12-037.

6. Alexiadis, V., & Chu, A. (2016). *Integrated Corridor Management Analysis, Modeling, and Simulation for the U.S.-75 Corridor in Dallas, Texas-Post-Deployment Assessment Report.* FHWA-JPO-16-396.

7. Amir, E. A., Davis, K. L., Tadmor, M. D., Simonds, E. F., Levine, J. H., Bendall, S. C., & Pe'er, D. (2013). t-SNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. *Nature Biotechnology, 31*(6), 545.

8. Barfield, W., Conquest, L., Spyridakis, J., & Haselkorn, M. (1989). Information Requirements for Real-Time Motorist Information Systems. *Vehicle Navigation and Information Systems Conference (VNIS)* (pp. 101-104). New York: IEEE.

9. Bekiaris-Liberis, N., Roncoli, C., & Papageorgiou, M. (2016). Highway traffic state estimation with mixed connected and conventional vehicles. *IEEE Transactions on Intelligent Transportation Systems, 17*(12), 3484-3497.

10. Ben-Akiva, M., Bierlaire, M., Koutsopoulos, H. N., & Mishalani, R. (2002). Real Time Simulation of Traffic Demand-Supply Interactions within DynaMIT. *Transportation and Network Analysis: Current Trends* (pp. 19-36). Boston, MA: Springer.

11. Benson, B. G. (1996). Motorist Attitudes about Content of Variable-Message Signs. *Transportation Research Record: Journal of the Transportation Research Board, 1550*, 48-57.

12. Box, D. (1998). *Essential COM.* Addison-Wesley Professional.

13. Branke, J., Goldate, P., & Prothmann, H. (2007). Actuated Traffic Signal Optimization using Evolutionary Algorithms. In Proceedings of the 6th European Congress and Exhibition on Intelligent Transport Systems and Services, 203-225.

14. Burvall, B., & Olegard, J. (2015). *A comparison of a genetic algorithm and simulated annealing applied to a traffic light control problem: A traffic intersection optimization problem.* KTH Royal Institute of Technology.

15. Campbell, R., & Skabardonis, A. (2014). Issues Affecting Performance of Adaptive Traffic Control Systems in Oversaturated Conditions. *Transportation Research Record: Journal of the Transportation Research Board, 2438*(1), 23-32.

16. Cao, X., Jiao, J., Zhang, Y., & Wang, X. (2019). Left-Turn Spillback Probability Estimation in a Connected Vehicle Environment. *Transportation Research Record*.

17. Ceylan, H. (2013). Optimal design of signal controlled road networks using differential evolution optimization algorithm. *Mathematical Problems in Engineering*.

18. Chatterjee, K. N., Hounsell, N. B., Firmin, P. E., & Bonsall, P. W. (2002). Bonsall. Driver Response to Variable Message Sign Information in London. *Transportation Research C, 10*(2), 149-169.

19. Chatterjee, k., & Mcdonald, M. (2004). Effectiveness of using variable message signs to disseminate dynamic traffic information: Evidence from field trails in European cities. *Transport Reviews, 24*(5), 559-585.

20. Chen, H., Rakha, H. A., & McGhee, C. C. (2013). Dynamic travel time prediction using pattern recognition. *20th World Congress on Intelligent Transportation Systems.* TU Delft.

21. Chen, J., & Xu, L. (2006). Road-junction traffic signal timing optimization by an adaptive particle swarm algorithm. *9th International Conference on Control, Automation, Robotics and Vision* (pp. 1-7). IEEE.

22. Chen, X., Qian, D., & Shi, D. (2011). Multi-objective Optimization Method of Signal Timing for the Non-motorized Transport at Intersection. *Journal of Transportation Systems Engineering and Information Technology, 11*(2), 106-111.

23. Chien, S. I., & Kuchipudi, C. M. (2003). Dynamic travel time prediction with real-time and historic data. *Journal of transportation engineering, 129*(6), 608-616.

24. Chilukuri, B. R., Perrin, J. J., & Martin, P. T. (2004). SCOOT and Incidents: Performance Evaluation in a Simulated Environment. *Transportation Research Record: Journal of Transportation Research Board, 1867*, 224-232.

25. Christofa, E., Ampountolas, K., & Skabardonis, A. (2016). Arterial traffic signal optimization: A person-based approach. *Transportation Research Part C: Emerging Technologies, 66*, 27-47.

26. Christofa, E., Argote, J., & Skabardonis, A. (2013). Arterial queue spillback detection and signal control based on connected vehicle technology. *Transportation Research Board: Journal of the Transportation Research Board, 2366*(1), 61-70.

27. Coello, C. A., Lamont, G. B., & Van Veldhuizen, D. A. (2007). *Evolutionary Algorithms for Solving Multi-Objective Problems* (Vol. 5). New York: Springer.

28. Dakic, I., Stevanovic, A., Zlatkovic, M., & Kergaye, C. (2017). Refinement of Performance Measures based on High-Resolution Signal and Detection Data. *Transportation Research Procedia, 22*, 372-381.

29. Day, C. M., & Bullock, D. M. (2011). Computational efficiency of alternative algorithms for arterial offset optimization. *Transportation Research Record: Journal of the Transportation Research Board, 2259*(1), 37-47.

30. Day, C. M., Brennan, T. M., Ernst, J. M., Sturdevant, J. R., & Bullock, D. M. (2011). *Data Reduction Procedures for Traffic Signal Systems Performance Measures.* Purdue University. Joint Transportation Research Program.

31. Day, C. M., Ernst, J. M., Brennan Jr, T. M., Chou, C. S., Hainen, A. M., Remias, S. M., . . . Bullock, D. M. (2012). Performance measures for adaptive signal control: Case study of system-in-the-loop simulation. *Transportation Research Record: Journal of the Transportation Research Board, 2311*(1), 1-15.

32. Day, C. M., Sturdevant, J. R., & Bullock, D. M. (2010). Outcome-oriented performance measures for management of signalized arterial capacity. *Transportation Research Record: Journal of the Transportation Research Board, 2192*(1), 24-36.

33. Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. A. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE transactions on evolutionary computation, 6*(2), 182-197.

34. Deeter, D. (2012). *Impacts of Traveler Information on the Overall Network.*

35. Deloche, F. (2019, March 16). *Recurrent Neural Network.* Retrieved from Wikipedia: https://en.wikipedia.org/wiki/Recurrent_neural_network#/media/File:Recurrent_neural_network_unfold.svg

36. Deloche, F. (2019, March 16). *Recurrent Neural Network.* Retrieved from Wikipedia: https://en.wikipedia.org/wiki/Recurrent_neural_network#/media/File:Long_Short-Term_Memory.svg

37. Dia, H., & Cottman, N. (2006). Evaluation of Arterial Incident Management Impacts Using Traffic Simulation. *IEEE Proceedings Intelligent Transportation Systems, 153(3)*, pp. 242-252.

38. Ding, C., and He, X. (2004). K-means Clustering via Principal Component Analysis. *Proceedings of the twenty-first international conference on Machine Learning* (p. 29). ACM.

39. Domenichini, L., Salerno, G., Fanfani, F., Bacchi, M., Giaccherini, A., & Costalli, L. B. (2012). Travel time in case of accident prediction model. *Social and Behavioral Sciences, 53*, 1080-1089.

40. Duan, Y., Lv, Y., & Wang, F. Y. (2016). Travel Time Prediction with LSTM Neural Network. *International Conference on Intelligent Transportation Systems (ITSC).* Brazil: IEEE.

41. Dunn Engineering Associates. (2006). *Alternate Route Handbook.* FHWA-HOP-06-092.

42. Dunteman, G. H. (1989). *Principal Components Analysis.* Sage.

43. Ezzat, A. A., Farouk, H. A., El-Kilany, K. S., & Moneim, A. F. (2014). Optimization using simulation of traffic light signal timings. *In Proceedings of the 2014 International Conference on Industrial Engineering and Operations Management.* Bali, Indonesia.

44. Fakharian Qom, S. (2016). *Multi-Resolution Modeling of Managed Lanes with Consideration of Autonomous/Connected Vehicle*. Miami: FIU Electronic Thesis and Dissertation.

45. FHWA. (2000). *Incident Management Successful Practices: A Cross-Cutting Study*. Washington, DC.: Federal Highway Administration.

46. FHWA. (2004). *Traffic analysis toolbox, Volume III: Guidelines for Applying Traffic Microsimulation Modelling Software.* McLean, VA: Federal Highway Administration, U.S. Department of Transportation.

47. FHWA. (2009). *Manual on uniform traffic control devices (MUTCD).*

48. FHWA. (2009). *Traffic Incident Management (TIM) Performance Measurement: On the Road to Success.* Federal Highway Administration, Office of Operations.

49. FHWA. (2012a). *Analysis, Modeling, and Simulation for Traffic Incident Management Applications.* Federal Highway Administration, U.S. Department of Transportation.

50. FHWA. (2012b). *Traffic Analysis Toolbox Volume XIV: Guidebook on the Utilization of Dynamic Traffic Assignment in Modeling.* Washington, D.C: Federal Highway Administration.

51. FHWA. (2017, December 1). *Office of Highway Policy Information*. Retrieved from Federal Highway Administration: https://www.fhwa.dot.gov/policyinformation/tables/02.cfm

52. FHWA. (2021, February 15). *Traffic Signal Timing Manual.* Retrieved from U.S. Department of Transportation: https://ops.fhwa.dot.gov/publications/fhwahop08024/chapter6.htm

53. Foo, S., & Abdulhai, B. (2006). Evaluating the impacts of changeable message signs on traffic diversion. *Intelligent Transportation Systems Conference* (pp. 891-896). IEEE.

54. Foy, M. D., Benekohal, R. F., & Goldberg, D. E. (1992). Signal timing determination using genetic algorithms. *Transportation Research Record, 1365,* 108.

55. Gers, F. A., Schraudolph, N. N., & Schmidhuber, J. (2003). Learning Precise Timing with LSTM Recurrent Networks. *Journal of Machine Learning Research, 3,* 115-143.

56. Giles, C. L., Kuhn, G. M., & Williams, R. J. (1994). Dynamic Recurrent Neural Networks: Theory and Applications. *IEEE Transactions on Neural Networks, 5*(2), 153-156.

57. Girianna, M., & Benekohal, R. F. (2002). Dynamic Signal Coordination for Networks with Oversaturated Intersections. *Transportation Research Record: Journal of the Transportation Research Board, 1811*, 122-130.

58. Girianna, M., & Benekohal, R. F. (2004). Using Genetic Algorithms to Design Signal Coordination for Oversaturated Networks. *Journal of Intelligent Transportation Systems, 8*, 117-129.

59. Goldberg, D. E., & Holland, J. H. (1988). Genetic algorithms and machine learning. *Machine Learning, 3*(2), 95-99.

60. Goodall, N. J., Smith, B. L., & Park, B. B. (2016). Microscopic estimation of freeway vehicle positions from the behavior of connected vehicles. *Journal of Intelligent Transportation Systems, 20*(1), 45-54.

61. Graves, A. (2012). *Supervised Sequence Labelling.* Springer.

62. Graves, A., & Schmidhuber, J. (2005). Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Networks, 18*(5-6), 602-610.

63. Guler, S. I., Menendez, M., & Meier, L. (2014). Using connected vehicle technology to improve the efficiency of intersections. *Transportation Research Part C: Emerging Technologies, 46*, 121-131.

64. Hadi, M., & Wallace, C. E. (1993). Hybrid Genetic Algorithm to Optimize Signal Phasing and Timing. *Transportation Research Record: Journal of Transportation Research Board, 1421*, 104-112.

65. Hadi, M., & Wallace, C. E. (1994). Optimization of signal phasing and timing using Cauchy simulated annealing. *Transportation Research Record: Journal of the Transportation Research Board*.

66. Hadi, M., Xiao, Y., & Rojas, M. (2013). Estimation of Diversion Rate During Incidents on Basis of Main-Line Detector Data. *Transportation Research Record: Journal of the Transportation Research Board, 2396*, 54-60.

67. Haghani, A., Hamedi, M., Fish, R. L., & Nouruzi, A. (2013). *Evaluation of dynamic message signs and their potential impact on traffic flow.*

68. Hajbabaie, A., & Benekohal, R. F. (2013). Traffic signal timing optimization: Choosing the objective function. *Transportation Research Record: Journal of the Transportation Research Board, 2355*(1), 10-19.

69. Hartigan, J. A., and Wong, M. A. (1979). A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics), 28(1)*, 100-108.

70. Haupt, R. L., & Haupt, S. E. (2004). *Practical Genetic Algorithms* (2nd Edition ed.). Wiley.

71. HCM. (2016). *Highway Capacity Manual.* Transportation Research Board.

72. He, D., Ni, F., & Yang, S. (2010). Prediction of Freeway Travel Time in Incident Management Evaluation Based on Genetic Neural Network. *Seventh International Conference on Traffic and Transportation StudiesAmerican Society of Civil EngineersSystems.*

73. He, J., & Hou, Z. (2012). Ant colony algorithm for traffic signal timing optimization. *Advances in Engineering Software, 43*(1), 14-18.

74. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735-1780.

75. Hochreiter, S., Bengio, Y., Frasconi, P., & Schmidhuber, J. (2001). *Gradient flow in recurrent nets: The difficulty of learning long-term dependencies.* IEEE Press.

76. Hojati, A. T., Ferreira, L., Washington, S., Charles, P., & Shobeirinejad, A. (2016, April). Modeling the Impact of Traffic Incidents on Travel Time Reliability. *Transportation Research Part C: Emerging Technologies, 65*, 49-60.

77. Iqbal, S. M., Khazraeian, S., & Hadi, M. (2018). A Methodology to Assess the Quality of Travel Time Estimation and Incident Detection based on Connected Vehicle Data. *Transportation Research Record, 2672*(42), 203-212.

78. Jacob, C., & Abdulhai, B. (2010). Machine learning for multi-jurisdictional optimal traffic corridor control. *Transportation Research Part A: Policy and Practice, 44*(2), 53-64.

79. Jain, A. K., & Dubes, R. C. (1988). *Algorithms for Clustering Data*. Prentice-Hall, Inc.

80. Jianming, H., Qiang, M., Qi, W., Jiajie, Z., & Yi, Z. (2012). Traffic congestion identification based on image processing. *IET Intelligent Transport Systems, 6*(2), 153-160.

81. Kesur, K. B. (2010). Generating More Equitable Traffic Signal Timing Plans. *Transportation Research Record: Journal of Transportation Research Board, 2192*(1), 108-115.

82. Ketchen, D. J., & Shook, C. L. (1996). The Application of Cluster Analysis in Strategic Management Research: An Analysis and Critique. *Strategic Management Journal, 17*(6), 441-458.

83. Khattak, A. J., Koppelman, F., & L, S. J. (1993). Stated preferences for investigating commuters' diversion propensity. *Transportation, 20*, 107-127.

84. Khattak, A. J., Wang, X., Zhang, H., & Cetin, M. (2011). *Primary and secondary incident management: predicting durations in real-time.* Virginia Center for Transportation Innovation and Research.

85. Khattak, A., Al-Deek, H., Yim, Y., & Hall, R. (1992). *Bay Area ATIS testbed plan.* Institute of Transportation Studies, University of California at Berkeley, California.

86. Kim, J., Mahmassani, H. S., Hou, T., & Alfelor, R. M. (2014). Development of Real-Time Simulation Based Decision Support System for Weather Responsive Traffic

Signal Operations. *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)* (pp. 810-815). IEEE.

87. Kim, W., Kim, H., Won, M., & Chang, G. L. (2017). *Development of a Traffic Management Decision Support Tool for Freeway Incident Traffic Management (FITM) Plan Deployment.* Maryland: State Highway Administration.

88. Knoop, V. L., Hoogendoorn, S. P., & Van Zuylen, H. J. (2008). Capacity reduction at incidents: empirical data collected from a helicopter. *Transportation Research Record, 2071*(1), 19-25.

89. Krishnakumari, P., Nguyen, T., Heydenrijk-Ottens, L., Vu, H. L., & van Lint, H. (2017). Traffic congestion pattern classification using multiclass active shape models. *Transportation Research Record*, *2645*(1), 94-103.

90. Kurzhanskiy, A. A., & Varaiya, P. (2010). Active traffic management on road networks: a macroscopic approach. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 368*(1928), 4607-4626.

91. Lampinen, J. (2002). A constraint handling approach for the differential evolution algorithm. *Proceedings of the 2002 Congress on Evolutionary Computation* (pp. 1468-1473). IEEE.

92. Lee, J., Abdulhai, B., Shalaby, A., & Chung, E. H. (2005). Real-time optimization for adaptive traffic signal control using genetic algorithms. *Journal of Intelligent Transportation Systems, 9*(3), 111-122.

93. Li, Y., Yu, L., Tao, S., & Chen, K. (2013). Multi-objective optimization of traffic signal timing for oversaturated intersection. *Mathematical Problems in Engineering*.

94. Li, Z., & Schonfeld, P. (2015). Hybrid simulated annealing and genetic algorithm for optimizing arterial signal timings under oversaturated traffic conditions. *Journal of advanced transportation, 49*(1), 153-170.

95. Lidbe, A. D., Tedla, E. G., Hainen, A. M., & Jones Jr, S. L. (2019). Feasibility Assessment for Implementing Adaptive Traffic Signal Control. *Journal of Transportation Engineering, Part A: Systems, 145*(2).

96. Lin, P. S., & Kou, C. C. (2003). Is an Alternative Route More Effective During a Freeway Incident? *ITE 2003 Annual Meeting and Exhibit Compendium of Technical Papers.* Washington DC: Institute of Transportation Engineers.

97. Liu, H. X., & Hu, H. (2013). *Improving Traffic Signal Operations for Integrated Corridor.* MnDOT.

98. Liu, H., & Ma, W. (2008). Real-time performance measurement system for arterial traffic signals. *Transportation Research Record: Journal of the Transportation Research Record*.

99. Liu, H., Van Zuylen, H., Van Lint, H., & Salomons, M. (2006). Predicting urban arterial travel time with state-space neural networks and Kalman filters. *Transportation Research Record: Journal of the Transportation Research Board, 1968*(1), 99-108.

100. Liu, Y., Chang, G.-L., & Yu, J. (2011). An Integrated Control Model for Freeway Corridor Under Nonrecurrent Congestion. *IEEE Transactions on Vehicular Technology, 60*(4), 1404-1418.

101. Liu, Y., Hu, Y., & Li, X. (2015). *A Smartphone-Based Prototype System for Incident/Work Zone Management Driven by Crowd-Sourced Data.* Interns Project Reports.

102. Liu, Y., Wang, Y., Yang, X., & Zhang, L. (2017). Short-term travel time prediction by deep learning: A comparison of different LSTM-DNN models. *IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)* (pp. 1-8). IEEE.

103. Lo, H. K., & Chow, A. H. (2004). Control strategies for oversaturated traffic. *Journal of Transportation Engineering, 130*(4), 466-478.

104. Luo, Z., & Liu, Y. (2011). Detour or No Detour? A New Decision Support Tool for Urban Freeway Incident Management. *2011 Transport Chicago conference.*

105. Ma, X., Tao, Z., Wang, Y., Yu, H., & Wang, Y. (2015). Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transportation Research Part C: Emerging Technologies, 54*, 187-197.

106. Madanat, S. M., Yang, C. Y., & Yen, Y. M. (1995). Analysis of Stated Route Diversion Intentions under Advanced Traveler Information Systems Using Latent Variable Modeling. *Transportation Research Record: Journal of the Transportation Research Board, 1485*, 10-17.

107. Mahapatra, D., Bozorgtabar, B., & Shao, L. (2020). Pathological Retinal Region Segmentation from OCT Images using Geometric Relation based Augmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 9611-9620.

108. Mahmassani, H. (2001). Dynamic Network Traffic Assignment and Simulation Methodology for Advanced System Management Applications. *Networks and Spatial Economics, 1*, 267-292.

109. Martin, P. T., Chaudhuri, P., Tasic, I., & Zlatkovic, M. (2011). *Freeway Incident: Simulation and Analysis.* Civil and Environmental Engineering, University of Utah.

110. Masghati-Amoli, G., Zarin, B., Farokhi, S. K., & Franz, M. L. (2015). Incident Capacity Reduction on Four-lane Freeways Using Real-World Data: Multilane Closures. *Transportation Research Board 94th Annual Meeting.* TRB.

111. Massahi, A., Hadi, M., Shams, K., & Baqersad, M. (2019). Evaluating Incident Responsive Signal Control Plans using Multi-Resolution Modeling. *Transportation Research Record: Journal of Transportation Research Board, 2673*(10), 804-813.

112. Matten. L., & Hinton, G. (2008). Visualizing Data Using t-SNE. *Journal of Machine Learning Research, 9,* 2579-2605.

113. Mazzenga, N. J., & Demetsky, M. J. (2009). *Investigation of Solutions to Recurring Congestion on Freeways.* Virginia Transportation Research Council.

114. Memon, G. Q., & Bullen, A. G. (1996). Multivariate optimization strategies for real-time traffic control signals. *Transportation Research Record: Journal of the Transportation Research Board, 1554*(1), 36-42.

115. Meng, S., Fu, Y., Liu, T., and Li, Y. (2015). Principal Component Analysis for Clustering Temporomandibular Joint Data. *Computational Intelligence and Design (ISCID), 8th International Symposium* (pp. 422-425). IEEE.

116. Miller, A. J. (1963). Setting for Fixed-Cycle Traffic Signals. *Oper. Res. Q., 14*(4), 373-386.

117. Mirchandani, P., Zhou, X., & Liu, J. (2018). *Developing a Multi-Resolution Traffic Simulation Platform for Integrated Active Traffic Operations Evaluation in Metropolitan Areas.* Maryland: National Transportation Center.

118. Mitchell, M. (1995). Genetic algorithms: An overview. *Complexity, 1*(1), 31-39.

119. Morshed, S. A., Arafat, M., Ashraf, A. M., Saha, R. (2020). Discovering the Commuters' Assessments on Disaster Resilience of Transportation Infrastructure. In International Conference on Transportation and Development 2020, (23-34). Reston, VA: American Society of Civil Engineers.

120. Mwakalonge, J. L., Mamun, M. M. H., Hong, J. D. & Chowdhury, M. R. (2019). Infrastructure and Policy Needs for Personal Electric Mobility Devices in a Connected Vehicle World.

121. Nafis, S., Alluri, P., Wu, W., & Kibria, B .M. G. (2021). Wrong-way driving crash injury analysis on arterial road networks using parametric data mining techniques. *Journal of Transportation Safety and Security.*

122. Neudorff, L. G., Randall, J., Reiss, R. A., & Gordon, R. L. (2003). *Freeway management and operations handbook.* FHWA.

123. Ng, L., Barfield, W., & Mannering, F. (1995). A survey-based methodology to determine information requirements for advanced traveler information systems. *Transportation Research Part C: Emerging Technologies, 3*(2), 113-127.

124. Noguchi, H., Maeda, T., Tamai, N., Minematsu, T., Sanada, H., & Mori, T. (2019). Pattern detection from seating pressure distribution during wheelchair motion using deep embedded clustering. *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 908-911). IEEE.

125. Ozbay, K., Xiao, W., Jaiswal, G., Bartin, B., Kachroo, P., & Baykal-Gursoy, M. (2009). Evaluation of incident management strategies and technologies using an integrated traffic/incident management simulation. *World Review of Intermodal Transportation Research*, 155-186.

126. Pan, B., Demiryurek, U., Shahabi, C., & Gupta, C. (2013). Forecasting spatiotemporal impact of traffic incidents on road networks. *2013 IEEE 13th International Conference on Data Mining* (pp. 587-596). IEEE.

127. Park, B. B., Rouphail, N. M., & Sacks, J. (2001). Assessment of stochastic signal optimization method using microsimulation. *Transportation Research Record: Journal of the Transportation Research Board, 1748*(1), 40-45.

128. Park, B., Messer, C. J., & Urbanik, T. (1999). Traffic signal optimization program for oversaturated conditions: genetic algorithm approach. *Transportation Research Record: Journal of the Transportation Research Board*, *1683*(1), 133-142.

129. Plaisant, C., Tarnoff, P., Saraf, A., & Rose, A. (1998). *Understanding Transportation Management Systems Performace with a Simulation-Based Learning Environment.* University of Maryland.

130. PTV AG. (2019). *PTV Vissim 11 User Manual.* Karlsruhe, Germany: PTV AG.

131. PTV AG. (2019). *PTV Visum 2020 - Manual.* Karlsruhe, Germany: PTV AG.

132. Putha, R., & Quardriifoglio, L. (2010). Using Ant Colony Optimization for Solving Traffic Signal Coordination in Oversaturated Networks. *89th Annual Meeting of the Transportation Research Board.* Washington, D.C.

133. Putha, R., Quadrifoglio, L., & Zechman, E. (2012). Comparing Ant Colony Optimization and Genetic Algorithm Approaches for Solving Traffic Signal Coordination under Oversaturation conditions. *Computer-Aided Civil and Infrastructure Engineering*, *27*(1), 14-28.

134. Robinson, E., Jacobs, T., Frankle, K., Serulle, N., & Pack, M. (2012). *Deployment, Use, and Effect of Real-Time Traveler Information Systems.*

135. Roess, R. P., McShane, W. R., & Prassas, E. S. (2013). *Traffic Engineering.* Pearson.

136. Rouphail, N. M., Park, B. B., & Sacks, J. (2000). Direct signal timing optimization: Strategy development and results. *XI Pan American Conference in Traffic and Transportation Engineering.*

137. Rousseeuw, P. J. (1987). Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis. *Journal of Computational and Applied Mathematics*, 53-65.

138. Saha, R., Tariq, M. T., Hadi, M., Xiao, Y. (2019). Pattern Recognition using Clustering Analysis to support transportation system management, Operations, and Modeling. *Journal of Advance Transportation.*

139. Sandt, A., Al-Deek, H., Rogers Jr, J. H., & Kayes, M. I. (2017). Using Agency Surveys and Benefit–Cost Analysis to Evaluate Highway Advisory Radio as Regional Traveler Information and Communication Tool. *Transportation Research Record, 2616*(1), 81-90.

140. Shen, L. (2008). *Freeway travel time estimation and prediction using dynamic neural networks.* FIU Digital Commons.

141. Siripanpornchana, C., Panichpapiboon, S., & Chaovalit, P. (2016). Travel-time prediction with deep learning. *2016 IEEE Region 10 Conference (TENCON)* (pp. 1859-1862). IEEE.

142. Smaglik, E. J., Sharma, A., Bullock, D. M., Sturdevant, J. R., & Duncan, G. (2007). Event-based data collection for generating actuated controller performance measures. *Transportation Research Record: Journal of the Transportation Research Board, 2035*(1), 97-106.

143. Smith, B. L., Qin, L., & Venkatanarayana, R. (2003). Characterization of Freeway Capacity Reduction Resulting from Traffic Accidents. *Journal of Transportation Engineering, 129*, 362-368.

144. Stevanovic, A., Martin, P. T., & Stevanovic, J. (2007). Vissim-Based Genetic Algorithm Optimization of Signal Timings. *Transportation Research Record: Journal of the Transportation Research Board, 2035*, 59-68.

145. Stevanovic, A., Stevanovic, J., & Kergaye, C. (2013). Optimization of traffic signal timings based on surrogate measures of safety. *Transportation research part C: Emerging Technologies, 32*, 159-178.

146. Sturdevant, J. R., Overman, T., Raamot, E., Deer, R., Miller, D., Bullock, D. M., . . . Remias, S. M. (2012). *Indiana traffic signal hi-resolution data logger enumerations.*

147. Sun, D., Benekohal, R. F., & Waller, S. T. (2003). Multiobjective traffic signal timing optimization using non-dominated sorting genetic algorithm. *In IEEE IV2003 Intelligent Vehicles Symposium. Proceedings* (pp. 198-203). IEEE.

148. Tariq, M. T.,Hadi, M., & Saha, R.(2021). Using High-Resolution Controller Data in the Calibration of Traffic Simulation Models. *Transportation Research Record: Journal of the Transportation Research Board*, 2021.

149. Tariq, M. T., Massahi, A., Saha, R., & Hadi, M. (2020). Combining Machine Learning and Fuzzy Rule-Based System in Automating Signal Timing Experts' Decisions during Non-Recurrent Congestion. *Transportation Research Record: Journal of Transportation Research Board, 2674*(6), 163-176.

150. Tariq, M. T., Saha, R. C., & Hadi, M. (2019). Methodology to Derive Route Diversion During Freeway Incident Conditions Based on Field Data. In 98th Transportation Research Board Annual Meeting. Washington D.C.

151. Tatomir, B., Rothkrantz, L. J., & Suson, A. C. (2009). Travel time prediction for dynamic routing using ant-based control. *Winter simulation conference* (pp. 1069-1078). Winter Simulation Conference.

152. Taylor, W. C., & Narupiti, S. (1996). *The Model Analysis Report on the Benefits of SCATS in Alleviating the Impacts of Incidents.* East Lansing, Michigan: Michigan State University.

153. Teklu, F., Sumalee, A., & Watling, D. (2007). A genetic algorithm approach for optimizing traffic control signals considering routing. *Computer-Aided Civil and Infrastructure Engineering, 22*(1), 31-43.

154. Texas Transportation Institute. (2003). *The 2003 Annual Urban Mobility Report.* Texas, U.S.A.: Texas A&M University System.

155. Thianniwet, T., Phosaard, S., & Pattara-Atikom, W. (2009). Classification of road traffic congestion levels from GPS data using a decision tree algorithm and sliding windows. *Proceedings of the world congress on engineering*, (pp. 1-3).

156. Tian, Z. Z., Balke, K., Engelbrecht, R., & Rilet, L. (2002). Integrated Control Strategies for Surface Street and Freeway Systems. *Transportation Research Record: The Journal of the Transportation Research Board*, 92-99.

157. Tiaprasert, K., Zhang, Y., Wang, X. B., & Zeng, X. (2015). Queue length estimation using connected vehicle technology for adaptive signal control. *IEEE Transactions on Intelligent Transportation Systems, 16*(4), 2129-2140.

158. Urbanik, T., Humphreys, D., Smith, B., & Levine, S. (2006). *Coordinated Freeway And Arterial Operations Handbook.* FHWA. FHWA.

159. Wang, Y., Cheevarunothai, P., & Hallenbeck, M. (2008). *Quantifying Incident-Induced Travel Delays on Freeways Using Traffic Sensor Data.* Transportation Northwest (Organization).

160. Webster, F. V. (1958). *Traffic Signal Setting.* Road Res. Lab., HMSO, London, U.K., Tech.

161. Wirtz, J. J., Schofer, J. L., & Schulz, D. F. (2005). Using simulation to test traffic incident management strategies: The benefits of preplanning. *Transportation Research Record: Journal of the Transportation Research Board, 1923*(1), 82-90.

162. Wright, B., Zou, Y., & Wang, Y. (2015). Impact of Traffic Incidents on Reliability of Freeway Travel Times. *Transportation Research Record: Journal of the Transportation Research Board, 2484*(1), 90-98.

163. Xia, J., Chen, M., & Qian, Z. (2010). Predicting Freeway Travel Time Under Incident Conditions. *Transportation Research Record: Journal of The Transportation Research Board*, 58-66.

164. Xie, J., Girshick, R., & Farhadi, A. (2016). Unsupervised deep embedding for clustering analysis. *International conference on machine learning*, (pp. 478-487).

165. Yang, S., & Qian, S. (2019). Understanding and predicting travel time with spatio-temporal features of network traffic flow, weather and incidents. *arXiv preprint arXiv:1901.06766*.

166. Yang, X., Lu, Y., & Chang, G. L. (2014). Dynamic signal priority control strategy to mitigate off-ramp queue spillback to freeway mainline segment. *Transportation Research Record, 2438*(1), 1-11.

167. Yazici, A., Kamga, C., Mudigonda, S., & Almotahari, S. (2018). *Reducing Incident-Induced Emissions and Energy Use in Transportation: Use of Social Media Feeds as an Incident Management Support Tool.* NYSERDA.

168. Yu, J., Park, S., & Chang, G. L. (2010). Advanced traveler information system for guiding route choice to ocean city, Maryland. *Transportation Research Record: Journal of the Transportation Research Board*, *2189*, 56-67.

169. Zhan, C., Gan, A., & Hadi, M. (2009). Identifying secondary crashes and their contributing factors. *Transportation research record*, *2102*(1), 68-75.

170. Zhang, L., Yin, Y., & Lou, Y. (2010). Robust signal timing for arterials under day-to-day demand variations. *Transportation Research Record: Journal of the Transportation Research Record*, *2192*(1), 156-166.

171. Zhang, Z., He, Q., Tong, H., Gou, J., & Li, X. (2016). Spatial-temporal traffic flow pattern identification and anomaly detection with dictionary-based compression theory in a large-scale urban network. *Transportation Research Part C: Emerging Technologies, 71*, 284-302.

172. Zhao, H., He, R., & Su, J. (2018). Multi-Objective Optimization of Traffic Signal Timing Using Non-Dominated Sorting Artificial Bee Colony Algorithm for Unsaturated Intersections. *Archives of Transport*.

173. Zhao, J., Gao, Y., Tang, J., Zhu, L., & Ma, J. (2018). Highway Travel Time Prediction Using Sparse Tensor Completion Tactics and-Nearest Neighbor Pattern Matching Method. *Journal of Advanced Transportation*.

174. Zheng, J., & Liu, H. X. (2017). Estimating traffic volumes for signalized intersections using connected vehicle data. *Transportation Research Part C: Emerging Technologies*, *79*, 347-362.

175. Zhou, H. (2008). Evaluation of route diversion strategies using computer simulation. *Journal of Transportation Systems Engineering and Information Technology, 8*(1), 61-67.

176. Zhu, Y., Gao, N., Wang, J., & Liu, C. (2016). Study on traffic flow patterns identification of single intersection intelligent signal control. *Procedia Engineering, 137*, 452-460.

177. Zlatkovic, M., & Zhou, X. (2015). Integration of Signal Timing Estimation Model and Dynamic Traffic Assignment in Feedback Loops: System Design and Case Study. *Journal of Advanced Transportation, 49*(6), 683-699.

VITA

RAJIB CHANDRA SAHA

EDUCATION

2004 – 2009	B.S., Civil Engineering Bangladesh University of Engineering & Technology, Dhaka, Bangladesh
2011 – 2013	MDS, Development Studies University of Dhaka, Dhaka, Bangladesh
2015 – 2017	M.S., Civil Engineering University of North Dakota, Grand Forks, North Dakota
2017 – 2021	Doctoral Candidate, Civil Engineering Department of Civil and Environmental Engineering, Florida International University, Miami, Florida
2018	Annual Book Scholarship, GCCITE
2019	ITE Florida District 10 Henry P. Boggs Student Paper Award
2020	Lifesavers Traffic Safety Scholar Award
2020 – 2021	Dissertation Year Fellowship, University Graduate School, Florida International University, Miami, Florida
2017 – 2018	Member Secretary, ITE FIU Student Chapter
2017 – 2018	CSO Representative, WTS FIU Student Chapter
2018 – 2019	Vice President, ITE FIU Student Chapter
2018 – 2019	Vice President, WTS FIU Student Chapter
2020	Chair, Student Leadership Summit

PUBLICATIONS AND PRESENTATIONS

1. Saha, R., T.M. Tariq, M. Hadi, and Y. Xiao. Pattern Recognition using Clustering Analysis to Support Transportation System Management, Operations. *Journal of Advanced Transportation*, 2019.

2. Tariq, T. M., R. Saha, and M. Hadi. Methodology to Derive Route Diversion During Freeway Incident Conditions Based on Field Data. In 98th Transportation Research Board Annual Meeting, 2019.

3. Hadi, M., Y. Xiao, M. S. Iqbal, T. Wang, R. Saha and M.T. Tariq. Data and Modeling Support of Off-Line and Real-Time Decisions Associated with Integrated Corridor Management. Florida International University, 2019.

4. Saha, R., T. M. Tariq, and M. Hadi. Deep Learning Approach for Predictive Analytics to Support Diversion during Freeway Incidents. *Transportation Research Record: Journal of the Transportation Research Board*, 2020.

5. Tariq, T. M., A. Massahi, R. Saha, and M. Hadi. Combining Machine Learning and Fuzzy Rule-Based System in Automating Signal Timing Experts Decisions during Non-Recurrent Congestion. *Transportation Research Record: Journal of the Transportation Research Board*, 2020.

6. Saha, R., T. M. Tariq, and M. Hadi. Identification of Traffic Congestion using Connected Vehicle for Active Traffic Management. In ASCE International Conference on Transportation & Development, 2020.

7. Tariq, T. M., M. Hadi, and R. Saha. Using High-Resolution Controller Data in the Calibration of Traffic Simulation Models. *Transportation Research Record: Journal of the Transportation Research Board*, 2021.

8. Saha, R., M. Hadi, T. Hunsanon, and T. M. Tariq. Integrated Approach for Optimizing Traffic Management Plans during Freeway Incidents. In 100th Transportation Research Board Annual Meeting, 2021.

9. Tariq, T. M., R. Saha, and M. Hadi. Sensor Data-Based Prediction of Route Diversion during Freeway Incidents, *Transportation Research Record: Journal of the Transportation Research Board*, 2021. (Under Review)