Predictive Traffic Operations and Control of Connected and Automated Vehicle Systems

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Civil and Environmental Engineering

By

Amr Elfar

EVANSTON, ILLINOIS

June 2019
ABSTRACT

Predictive Traffic Operations and Control of Connected and Automated Vehicle Systems

Connected and automated vehicle (CAV) technology is a disruptive transportation development with potentially transformative impacts on society and the economy. CAV systems promise to significantly reduce human-caused road crashes, improve traffic flow performance, and lower pollutant emissions. However, realizing those benefits requires strategic planning for the deployment of CAV systems and developing advanced traffic control algorithms that utilize their new capabilities. To that end, the objective of this dissertation is to develop innovative traffic management strategies that utilize the big stream of data generated by CAV systems and the predictive capability of machine learning algorithms.

The dissertation starts by introducing a methodological framework for developing predictive traffic management and control strategies utilizing CAV systems. It considers three main components: 1) traffic monitoring, 2) traffic state prediction, and 3) control strategy. Following this framework, the dissertation presents a novel method to identify shockwave formation and track its propagation based on the speed distribution of individual vehicles available through connected vehicles technology. The analysis shows a consistent pattern where shockwave formation, indicated by a speed drop propagating over space and time, is associated with a sharp increase in the value of speed standard deviation (SSD).
Building on the aforementioned method, the dissertation also presents online and offline models for short-term traffic congestion prediction. Offline models are calibrated based on historical data and are updated (re-trained) whenever significant changes occur in the system, such as changes to the infrastructure. Online models are calibrated using historical data and updated regularly using real-time information on prevailing traffic conditions broadcasted by CAVs. Results show that the accuracy of the proposed models can reach 97%.

Utilizing the early shockwave detection method and the congestion prediction models, the dissertation presents a predictive speed harmonization system with two CAV control strategies: centralized and decentralized. The centralized system relies on a traffic management center to collect data from CAVs within a road segment of interest, predict traffic congestion, and broadcast updated speed limits to CAVs in order to mitigate congestion. The decentralized system relies on individual CAVs to collect data through communicating with each other, predict traffic congestion using vehicle-specific models, and update their speed limits to mitigate congestion. Case studies of multiple operational scenarios show that both systems can reduce the severity and lengths of traffic shockwaves, improve the overall traffic stability, increase overall speed, and reduce travel time. The decentralized strategy can be more cost-efficient over the long run since it does not require any more resources beyond what CAVs are expected to have. For it to be effective, however, cross-communication between different CAV fleets is required.

Finally, the dissertation presents truck platooning as a special application of decentralized CAV traffic control strategies. Results show that forming platoons under an opportunistic strategy can be difficult due to the generally low number of trucks on highways.
ACKNOWLEDGEMENTS

I would like to express my sincerest appreciation to my advisor Prof. Hani Mahmassani for his support, mentorship, and guidance throughout my graduate career at Northwestern. He has been my role model of research excellence, dedication, and work ethics.

I would also like to extend my sincerest gratitude to Professor Joseph Schofer for his invaluable counsel and support throughout my graduate school journey. His dedication to teaching and mentoring the younger generation has always been an inspiration to me.

I would like to especially thank Dr. Alireza Talebpour for being an amazing mentor and a friend. I truly appreciate his help with my research and professional career.

I would also like to thank my friends and colleagues at Northwestern who made my graduate school journey an amazing one. Special thanks to Archak Mittal, Lama Bou Mjahed, Marija Ostojic, Mike Hyland, Haleh Ale-Ahmad, Lama Al Hajj Hassan, Eunhye Kim, Helen Pinto, Omer Verbas, and Ying Chen.

Finally, I would like to thank my dad, mom, sister, and brothers for their unconditional love and support and for being the bedrock that I can always fall back onto. Words cannot express how thankful I am to my mom for being my strongest supporter and for encouraging me to take risks and explore my interests. I dedicate this dissertation to her.
Dedication

To Mom
# TABLE OF CONTENTS

Abstract........................................................................................................................................... 3

List of Figures........................................................................................................................................ 13

List of Tables .......................................................................................................................................... 17

1. **Introduction** ................................................................................................................................... 18
   1.1 Problem Statement .......................................................................................................................... 18
   1.2 Research Objectives ....................................................................................................................... 19
   1.3 Research Contributions .................................................................................................................. 20
   1.4 Organization ......................................................................................................................................... 22

2. **Literature Review** ........................................................................................................................... 24
   2.1 Strategic and Operational Impacts of Connected and Automated Vehicle Systems ................. 24
       2.1.1 Demand Changes ......................................................................................................................... 27
       2.1.2 Supply Changes ............................................................................................................................ 29
       2.1.3 Operational Performance ............................................................................................................ 33
   2.2 Evaluating the Operational Impacts of Connected and Automated Vehicle Systems .......... 43
       2.2.1 Automated Intelligent Cruise Control/ Adaptive Cruise Control ............................................. 44
       2.2.2 Cooperative Adaptive Cruise Control .......................................................................................... 46
       2.2.3 Speed Harmonization in a Connected Vehicle Environment .................................................... 49
       2.2.4 Dedicated Lanes for Automated Vehicles ................................................................................ 52
       2.2.5 Queue Warning in a Connected Vehicle Environment ............................................................... 53
   2.3 Chapter Summary ............................................................................................................................ 56

3.1 Traffic Monitoring

3.1.1 Estimating Traffic Properties

3.1.2 Tracking Traffic Dynamics

3.2 Traffic State Prediction

3.2.1 Machine Learning Techniques

3.2.2 Prediction Model Type

3.3 Control Strategy

3.3.1 Centralized Traffic Control (V2I)

3.3.2 Decentralized CAV Traffic Control (V2V)

3.4 Chapter Summary

4. Traffic Shockwave Detection in a Connected Environment using the Speed Distribution of Individual Vehicles
4.3.7 Comparison between SSD Waves and Speed Wavelet Transformation for Detecting Shockwaves ................................................................. 88

4.4 Chapter Summary ............................................................................. 91


5.1 Data ..................................................................................................... 96

5.2 Methodology ......................................................................................... 97

5.2.1 Estimating Traffic Properties ............................................................ 97

5.2.2 Speed Standard Deviation as an Indicator for Congestion ................ 98

5.2.3 Identifying Traffic States using K-means Clustering of the Fundamental Diagram .......... 98

5.2.4 Temporally and Spatially Lagged Variables to Build Predictive Models ........... 100

5.2.5 Offline Predictive Models .................................................................. 103

5.2.6 Online (Real-time) Predictive Models .................................................. 104

5.2.7 Model Accuracy and K-fold Cross-Validation ....................................... 105

5.2.8 Predictive Models with Partial Connectivity .......................................... 107

5.2.9 Specifications of the Machine Learning Techniques .................................. 109

5.3 Results and Analysis ......................................................................... 109

5.3.1 Offline Predictive Models with Full Connectivity ................................. 109

5.3.2 Offline Predictive Models with Partial Connectivity .............................. 111

5.3.3 Online Predictive Models .................................................................... 113

5.4 Chapter Summary ............................................................................. 115

6. Centralized CAV Traffic Management Application – Predictive Speed Harmonization in a Connected Environment with Centralized Speed Control ................. 117
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Methodology</td>
<td>121</td>
</tr>
<tr>
<td>6.1.1</td>
<td>Traffic Monitoring</td>
<td>123</td>
</tr>
<tr>
<td>6.1.2</td>
<td>Congestion Prediction</td>
<td>124</td>
</tr>
<tr>
<td>6.1.3</td>
<td>Speed Control</td>
<td>126</td>
</tr>
<tr>
<td>6.2</td>
<td>CAV Traffic Microsimulation Tool</td>
<td>127</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Modeling Connected Vehicles</td>
<td>128</td>
</tr>
<tr>
<td>6.2.2</td>
<td>Modeling Automated Vehicles</td>
<td>129</td>
</tr>
<tr>
<td>6.3</td>
<td>Results and Analysis</td>
<td>131</td>
</tr>
<tr>
<td>6.3.1</td>
<td>Impact of the Predictive Speed Harmonization System on Traffic</td>
<td>133</td>
</tr>
<tr>
<td></td>
<td>Performance</td>
<td></td>
</tr>
<tr>
<td>6.3.2</td>
<td>System Performance in Partial Connectivity Conditions</td>
<td>138</td>
</tr>
<tr>
<td>6.3.3</td>
<td>System Performance in Mixed Traffic Conditions</td>
<td>140</td>
</tr>
<tr>
<td>6.4</td>
<td>Fine-tuning Design Parameters for Optimal Results</td>
<td>145</td>
</tr>
<tr>
<td>6.4.1</td>
<td>Optimization Formulation for the Speed Control Module</td>
<td>148</td>
</tr>
<tr>
<td>6.4.2</td>
<td>Performance Comparison: Optimization-based versus Decision-Tree</td>
<td>151</td>
</tr>
<tr>
<td></td>
<td>Speed Control</td>
<td></td>
</tr>
<tr>
<td>6.4.3</td>
<td>Applying optimization-based control in a real-world scenario</td>
<td>155</td>
</tr>
<tr>
<td>6.5</td>
<td>Chapter Summary</td>
<td>157</td>
</tr>
<tr>
<td>7.</td>
<td>Decentralized CAV Traffic Management Application – Predictive Speed</td>
<td>159</td>
</tr>
<tr>
<td></td>
<td>Harmonization in a Connected Environment with Decentralized Speed</td>
<td></td>
</tr>
<tr>
<td>7.1</td>
<td>Methodology</td>
<td>160</td>
</tr>
<tr>
<td>7.1.1</td>
<td>Traffic Monitoring</td>
<td>162</td>
</tr>
<tr>
<td>7.1.2</td>
<td>Congestion Prediction</td>
<td>163</td>
</tr>
<tr>
<td>7.1.3</td>
<td>Speed Control</td>
<td>164</td>
</tr>
</tbody>
</table>
7.2 Results and Analysis ........................................................................................................... 165

7.2.1 Impact of the Decentralized Predictive Speed Harmonization System on Traffic Performance .................................................................................................................. 166

7.2.2 System Performance in Partial Connectivity Conditions .............................................. 171

7.2.3 System Performance in Mixed Traffic Conditions ...................................................... 174

7.3 Centralized vs. Decentralized Speed Harmonization Systems ...................................... 178

7.4 Chapter Summary ............................................................................................................. 183

8. Decentralized CAV Traffic Management Application – Truck Platooning in Mixed Traffic Environment .......................................................................................................................... 185

8.1 Methodology .................................................................................................................... 186

8.1.1 Cruise Control (CC) ..................................................................................................... 186

8.1.2 Adaptive Cruise Control (ACC) .................................................................................. 187

8.1.3 Cooperative Adaptive Cruise Control (CACC) .......................................................... 187

8.1.4 Truck Platoon Formation ............................................................................................ 188

8.2 Results and Analysis ....................................................................................................... 188

8.2.1 Impact of Automated Truck Platooning in Mixed Traffic Scenarios – Low Automation (30 percent AV) Condition ........................................................................................................ 188

8.2.2 Impact of Automated Truck Platooning in Mixed Traffic Scenarios – High Traffic Automation (70 percent AV)........................................................................................................ 194

8.2.3 Truck Platoon Size Analysis ....................................................................................... 200

8.2.4 Truck Platoon Duration Analysis ................................................................................ 202

8.3 Chapter Summary ............................................................................................................. 203

9. Conclusion and Future Research ....................................................................................... 205

9.1 Conclusion ....................................................................................................................... 205
9.2 Future Research ........................................................................................................................................... 210

References .......................................................................................................................................................... 213
LIST OF FIGURES

FIGURE 1 A methodological framework for evaluating the strategic and operational impacts of CAV technology (1)..........................................................26

FIGURE 2 The demand changes component of the general CAV AMS framework (1).................................27

FIGURE 3 The supply changes component of the general CAV AMS framework (1)..........................30

FIGURE 4 The operational performance component of the general CAV AMS framework (1)........36

FIGURE 5 The network integration component of the general CAV AMS framework (1)........42

FIGURE 6 Queue Warning Concept ..........................................................54

FIGURE 7 Framework for developing predictive traffic control strategies of connected and automated vehicles ..........................................................60

FIGURE 8 Time-space diagrams for (a) mean speed 7:50AM – 8:05AM, (b) speed standard deviation 7:50AM – 8:05AM, (c) mean speed 8:05AM – 8:20AM, (d) speed standard deviation 8:05AM – 8:20AM, (d) mean speed 8:20AM – 8:35AM, (e) speed standard deviation 8:20AM – 8:35AM .........................77

FIGURE 9 Time-series graph for mean speed and speed standard deviation in section 4 during (a) period 7:50 - 8:05 and (b) period 8:05AM – 8:20AM ..........................................................78

FIGURE 10 Propagation of SSD spikes compared to the propagation of shockwaves during the period 7:50AM – 8:05AM. Direction of travel is from 3 – 7..................................................................80


FIGURE 12 Speed Standard Deviation waves vs. Speed shockwaves at different market penetrations (10%, 20%, 30%, 70%, and 100%) in section 2 during the period 7:50AM – 8:05AM .........................84

FIGURE 13 Number of vehicles used to estimate the speed standard deviation at different market penetrations for section 2 in period 7:50AM – 8:05AM.......................................................85

FIGURE 14 Propagation of SSD waves at partial market penetrations.....................................................87

FIGURE 15 Wavelet Transformation of Mean Speed vs. SSD Waves estimated for Section 3 During the Period 7:50AM – 8:05AM..............................................................89
FIGURE 16 Propagation of Wavelet Energy and SSD over Sections 3 – 6 during the Period 7:50AM – 8:05AM

FIGURE 17 Traffic states for the study segment of US-101

FIGURE 18 Model updating process for online predictive models using NGSIM data

FIGURE 19 Mean Speed and speed standard deviation estimated at different market penetrations for section 2 during the period 7:50AM – 8:05AM

FIGURE 20 Main components of a centralized predictive speed harmonization system

FIGURE 21 Speed decision tree for the SPDHRM system (61; 130; 131)

FIGURE 22 Maximum Safe Speed Curve

FIGURE 23 Two-lane highway segment for testing multiple SPDHRM operational scenarios, main lanes volume 3000 veh/hr, ramp volume 500 veh/hr

FIGURE 24 Impact of predictive SPDHRM on traffic shockwave formation (Centralized Control): (a) base case (b) active SPDHRM, broadcasting distance 1000m, prediction horizon 20s

FIGURE 25 Impact of predictive SPDHRM on traffic flow stability and breakdown (Centralized Control): (a) base case (b) active SPDHRM, broadcasting distance 1000m, prediction horizon 20s

FIGURE 26 Impact of predictive SPDHRM on segment speed distribution (Centralized Control): (a) base case (b) active SPDHRM, broadcasting distance 1000m, prediction horizon 20s

FIGURE 27 Impact of predictive SPDHRM on vehicle travel time (Centralized Control): (a) base case (b) active SPDHRM, broadcasting distance 1000m, prediction horizon 20s

FIGURE 28 Impact of partial connectivity on the SPDHRM performance (Centralized Control): a – c) flow density diagrams for 0%, 40%, and 80% CV market penetration respectively, d-f) travel time distributions for 0%, 40%, and 80% CV market penetration respectively, broadcasting distance 1000m, prediction horizon 20s

FIGURE 29 Effectiveness of the centralized SPDHRM system in mixed traffic environment – low automation: a, d) flow-density diagram and travel time distribution at 0% AV MPR with INACTIVE SPDHRM, b, e) flow-density diagram and travel time distribution at 30% AV MPR with INACTIVE SPDHRM, c, f) flow-density diagram and travel time distribution at 30% AV MPR with ACTIVE SPDHRM, broadcasting distance 1000m, prediction horizon 20s
FIGURE 30 Effectiveness of the centralized SPDHRM system in mixed traffic environment – high automation: a, d) flow-density diagram and travel time distribution at 0% AV MPR with INACTIVE SPDHRM, b, e) flow-density diagram and travel time distribution at 70% AV MPR with INACTIVE SPDHRM, c, f) flow-density diagram and travel time distribution at 70% AV MPR with ACTIVE SPDHRM, broadcasting distance 1000m, prediction horizon 20s................................. 144

FIGURE 31 Traffic shockwave formation patterns: a) speed decision-tree, b) optimization-based speed control ............................................................ 152

FIGURE 32 Flow-density diagrams: a) speed decision-tree, b) optimization-based speed control .......... 154

FIGURE 33 Speed distributions: a) speed decision-tree, b) optimization-based speed control ............. 154

FIGURE 34 Travel time distributions: a) speed decision-tree, b) optimization-based speed control ...... 154

FIGURE 35 Main components of a decentralized predictive speed harmonization system.............. 154

FIGURE 36 Impact of predictive SPDHRM on traffic shockwave formation (Decentralized Control): (a) base case (b) active SPDHRM, broadcasting distance 1000m, prediction horizon 20s, Single Prediction Model................................. 162

FIGURE 37 Impact of predictive SPDHRM on traffic flow stability and breakdown (Decentralized Control): (a) base case (b) active SPDHRM, broadcasting distance 1000m, prediction horizon 20s, Single Prediction Model........................................................................................................... 167

FIGURE 38 Impact of predictive SPDHRM on segment speed distribution (Decentralized Control): (a) base case (b) active SPDHRM, broadcasting distance 1000m, prediction horizon 20s, Single Prediction Model ......................................................................................................................... 168

FIGURE 39 Impact of predictive SPDHRM on vehicle travel time (Decentralized Control): (a) base case (b) active SPDHRM, broadcasting distance 1000m, prediction horizon 20s, Single Prediction Model... 170

FIGURE 40 Impact of partial connectivity on the SPDHRM’s performance (Decentralized Control): a – c) flow density diagrams for 0%, 40%, and 80% CV market penetration respectively, d-f) travel time distributions for 0%, 40%, and 80% CV market penetration respectively, broadcasting distance 1000m, prediction horizon 20s, single prediction model................................................................. 173

FIGURE 41 Impact of multiple CAV fleets on decentralized SPDHRM system: a, b) travel time distribution in main lanes at low connectivity level (40%) for multiple and single prediction models respectively, c, d) travel time distribution in main lanes at high connectivity level (80%) for multiple and single prediction models respectively, broadcasting distance 1000m, prediction horizon 20s ........................................... 174
FIGURE 42 Effectiveness of the centralized SPDHRM system in mixed traffic environment – low automation: a, d) flow-density diagram and travel time distribution at 0% AV MPR with INACTIVE SPDHRM, b, e) flow-density diagram and travel time distribution at 30% AV MPR with INACTIVE SPDHRM, c, f) flow-density diagram and travel time distribution at 30% AV MPR with ACTIVE SPDHRM, broadcasting distance 1000m, prediction horizon 20s.

FIGURE 43 Effectiveness of the decentralized SPDHRM system in mixed traffic environment – high automation: a, d) flow-density diagram and travel time distribution at 0% AV MPR with INACTIVE SPDHRM, b, e) flow-density diagram and travel time distribution at 70% AV MPR with INACTIVE SPDHRM, c, f) flow-density diagram and travel time distribution at 70% AV MPR with ACTIVE SPDHRM, broadcasting distance 1000m, prediction horizon 20s.

FIGURE 44 Fundamental diagrams for truck platooning scenarios at a low (30 percent) AV market penetration rate.

FIGURE 45 Overall travel time distribution for truck platooning scenarios at a low (30 percent) AV market penetration rate.

FIGURE 46 Truck travel time distribution for truck platooning scenarios at a low (30 percent) AV market penetration rate.

FIGURE 47 Car travel time distribution for truck platooning scenarios at a low (30 percent) AV market penetration rate.

FIGURE 48 Fundamental diagrams for truck platooning scenarios at a high (70 percent) AV market penetration.

FIGURE 49 Overall travel time distribution for truck platooning scenarios at a high (70 percent) AV market penetration rate.

FIGURE 50 Truck travel time distribution for truck platooning scenarios at a high (70 percent) AV market penetration rate.

FIGURE 51 Car travel time distribution for truck platooning scenarios at a high (70 percent) AV market penetration rate.

FIGURE 52 Truck platoon size for 10 percent/20 percent trucks at low (30 percent) and high (70 percent) AV market penetration levels.

FIGURE 53 Truck platoon duration for 10 percent/20 percent trucks at low (30 percent) and high (70 percent) AV market penetration levels.
LIST OF TABLES

TABLE 1 Summary of selected studies on the impacts of Automated Intelligent Cruise Control (AICC)/ Adaptive Cruise Control (ACC) and Cooperative Adaptive Cruise Control (CACC) on traffic flow ....... 44

TABLE 2 Summary of recent studies on speed harmonization applications enabled by CAV systems - Ma et al (65; 68) ........................................................................................................................................... 51

TABLE 3 Summary of the variables used in the predictive models .......................................................................................................................... 102

TABLE 4 Comparison of Offline Models with Full Connectivity ..................................................................................................................... 110

TABLE 5 Comparison of Offline Models with Partial Connectivity ........................................................................................................... 113

TABLE 6 Comparison of Online Predictive Model ............................................................................................................................................. 114

TABLE 7 summary of variables used in the predictive model ................................................................................................................................. 125

TABLE 8 Comparison of Traffic Congestion Prediction Models .................................................................................................................... 126

TABLE 9 Impact of Prediction Horizon on SPDHRM System Performance .................................................................................................... 146

TABLE 10 Impact of Broadcasting Distance on SPDHRM System Performance .......................................................................................... 147

TABLE 11 Control strategy comparison: decision-tree vs optimization based ............................................................................................... 157

TABLE 12 Centralized vs. Decentralized Speed Harmonization System ......................................................................................................... 183

TABLE 13 Automated truck platooning in mixed traffic scenarios – low (30 percent) AV market penetration condition (percent) .......................................................................................................................... 189

TABLE 14 Automated truck platooning in mixed traffic scenarios – high (70 percent) AV market penetration condition (percent) .......................................................................................................................... 194
1. INTRODUCTION

1.1 Problem Statement

Connected and automated vehicle (CAV) technology, vehicles with wireless telecommunication capabilities and partially or fully automated driving functions, is a disruptive transportation development with potentially transformative impacts on society and the economy (2; 3). CAV systems promise to significantly reduce human-caused road crashes, improve traffic flow performance, and lower pollutant emissions. However, realizing those benefits requires strategic planning for the deployment of CAV systems and developing advanced traffic control algorithms that utilize their new capabilities. Therefore, the motivation behind this dissertation is the design of new traffic control strategies that utilize the wireless telecommunication capabilities of CAVs and their new driving behavior to improve the overall performance of transportation systems.

A distinctive feature of CAV systems that is the focus of this research is their ability to generate and broadcast real-time information about their status through wireless telecommunications (4). Detailed information such as vehicle location, speed, and acceleration can be exchanged between the vehicles themselves (V2V communications) or between the vehicles and the infrastructure (V2I communications). The detailed trajectories can then be used to track finer patterns in vehicle movements and understand traffic flow dynamics to improve traffic prediction and control. Perturbations in traffic flow, for example, constitute one of the main factors for causing traffic congestion (5; 6) that could be identified using the V2V/V2I technology. In this
case, detailed vehicle trajectories could be used to measure disturbances caused by individual vehicles at a high resolution.

Utilizing the detailed vehicle information that would be generated by CAV systems is beyond the capability of current traffic controls strategies. Those rely on aggregated traffic properties, such as vehicles counts and mean speeds, which are typically collected by embedded road sensors. Therefore, developing effective traffic control and management strategies for CAV systems requires implementing new tools that can handle big streams of data such as machine learning. These techniques identify patterns within large data sets and use those patterns for prediction. In this case, machine learning algorithms would identify traffic patterns within the vehicle trajectories transmitted by CAV systems to predict traffic properties such as congestion.

1.2 Research Objectives

The main objective of this dissertation is to develop innovative traffic management strategies that utilize the big stream of data generated by CAVs and the predictive capability of machine learning algorithms. Those strategies will be tested for different operational scenarios in various traffic conditions. At the higher level, this research aims to:

1. Offer insights to operating agencies on potential strategies for utilizing data streams generated by CAV systems in traffic control
2. Develop a CAV microsimulation framework with machine learning capabilities to evaluate different strategies and operational scenarios
3. Extend the body of knowledge on traffic flow dynamics of CAV systems
1.3 Research Contributions

The contributions of this dissertation include:

- Introducing a methodological framework for developing predictive traffic management and control strategies of CAV systems; this allows:
  - Investigating the potential implications CAVs have on traffic monitoring, traffic state prediction, and control strategy

- Developing a novel method for identifying the formation and propagation of traffic shockwaves over small road sections using the speed distribution of individual vehicles in a connected environment; this allows:
  - Investigating the impacts of partial connectivity (percentage of connected vehicles) on shockwave identification and formation

- Building traffic congestion prediction models using three different machine learning techniques: logistic regression, random forests, and neural networks. This entails:
  - Introducing offline prediction models which are calibrated based on historical data and updated (re-trained) whenever significant changes occur in the system, such as changes/updates to the infrastructure
  - Introducing online models that are calibrated using historical data and updated regularly using real-time information on prevailing traffic conditions obtained through V2V/V2I communications
  - Examining the impact of partial connectivity (percentage of connected vehicles) on the prediction accuracy of the congestion models
• Developing predictive speed harmonization strategies for CAV systems that integrate the predictive capability of machine learning algorithms; this entails:
  o Introducing a centralized speed control strategy in which a central system continuously evaluates the state of the traffic and broadcasts updates speed limits to CAVs in order to mitigate traffic congestion
  o Introducing an optimization formulation for selecting the system’s design parameters (e.g. speed limit and broadcasting distance) such that it maximizes the overall traffic speed (minimizes travel time)
  o Introducing a decentralized speed control strategy in which CAVs individually adjust their speed based on the information they receive from a cluster or a fleet of vehicles in order to mitigate traffic congestion

• Evaluating the mobility impacts of the developed speed harmonization strategies in a mixed traffic environment; this includes:
  o Evaluating the impact of the different systems on traffic performance using measures such as shockwave formations, stability, speed, and travel time
  o Evaluating the impact of partial connectivity conditions on the performance of control strategies
  o Evaluating performance of the control strategies in mixed traffic conditions (i.e. with automated vehicles).

• Evaluating the impacts of an opportunistic truck platooning strategy in a mixed traffic environment as an example of decentralized traffic management strategies; this includes:
o Evaluating the impact of the strategy on traffic performance in low traffic automation conditions

o Evaluating the impact of the strategy in on traffic performance in high traffic automation conditions

1.4 Organization

This dissertation is organized as follows. Chapter 1 introduces the thesis and motivation of this dissertation. Chapter 2 provides a literature review of the strategic and operational impacts of CAV systems and traffic control strategies in a connected vehicle environment, setting the context in which the new CAV applications in this dissertation were developed. Chapter 1 introduces a methodological framework for developing traffic control and management strategies of CAV systems which serves as a road map for developing the predictive speed harmonization systems introduced in Chapters 6 and 7. Chapter 4 introduces a new shockwave detection method using the speed variation of vehicles, providing a proof of concept for potential traffic monitoring methods utilizing CAV trajectories. Chapter 5 introduces traffic congestion prediction models with machine learning algorithms, which provides a proof of concept for utilizing CAV data to predict traffic states. Building on the early shockwave detection method and congestion prediction models introduced in previous chapters, Chapter 6 introduces a predictive speed harmonization system with a centralized speed control strategy that relies on a central traffic management system to execute the control logic. Chapter 7 introduces a predictive speed harmonization system with a decentralized speed control strategy that relies on individual vehicles themselves to execute the control logic based on the information they share with each other. Chapter 8 presents an
opportunistic truck platooning strategy as a special application of decentralized traffic management strategies. Finally, Chapter 9 provides concluding remarks.
2. LITERATURE REVIEW

To provide the context in which the introduced CAV control applications were developed, this Chapter provides an overview of strategic and operational impacts of CAV systems. It also provides an overview of traffic control strategies that can be implemented in a connected vehicle environment. Those include cooperative adaptive cruise control, speed harmonization, queue warning, and dedicated lanes for automated vehicles. All of these strategies can potentially be improved by utilizing the predictive capability of machine learning algorithms as tested in later chapters in this dissertation.

2.1 Strategic and Operational Impacts of Connected and Automated Vehicle Systems

The implications of CAV technology are far reaching at the strategic and operational levels, yet those impacts are interdependent (2). On the strategic level, the technology will potentially affect the supply of mobility services, demand patterns, and travel behavior. On the operational/tactical level, the technology can potentially improve traffic flow performance of transportation facilities and networks.

CAV technology is expected to help introduce entirely new modes of mobility (7; 8) in the form of Shared-Automated-Vehicle (SAV) fleets, for example, in addition to improving multiple aspects of current mobility options. Such improvements include highly automating certain driving tasks, or all of them from origin to destination, and supporting travel-related decisions by providing real-time information through wireless telecommunications.
The availability of new mobility forms in addition to the improvements to current transportation systems through connectivity can affect the activity patterns (9-13) and mobility choices of travelers (14; 15). Those changes can involve household-level decision, such as owning a car, or individual decisions such as departure time and route choice (16).

Changes to both supply and demand in addition to the improvements brought by connectivity and automation to traffic flow ultimately affect the operational performance of transportation systems. The potential improvements include, for example, increased throughput (17-23) and improved safety through incorporating real-time information (24-27) on prevailing traffic conditions and the addition of a safer automated driving behavior (28-33).

FIGURE 1 illustrates the interactions among the supply, demand, and operation impact of CAV systems in a comprehensive framework (1). The framework includes four main components that needs to be integrated in a comprehensive CAV AMS system for an improved evaluation of CAV impacts. Those components are:

1. **Supply Changes**: to analyze the emergence of new mobility options enabled by CAVs and the changes incurred by the new technology to the infrastructure
2. **Demand Changes**: to evaluate CAV impacts on activity and travel choices
3. **Operational Performance**: to evaluate the impacts of the technology on the performance of transportation systems such as increased capacity and improved travel time
4. **Network Integration**: to capture the multi-agent interactions at the network level
FIGURE 1 A methodological framework for evaluating the strategic and operational impacts of CAV technology (1)
2.1.1 Demand Changes

The new forms of mobility (7; 8) enabled by CAV technology and their expected improvements to the performance of transportation systems could lead to fundamental changes to the transport-related decisions. Those changes could affect 1) the activity patterns (9-13) and 2) mobility choices of travelers (14; 15) at multiple levels, as illustrated in FIGURE 2.

FIGURE 2 The demand changes component of the general CAV AMS framework (1)

Activity patterns

On the higher level, the potentially improved features of the new mobility options can impact the activity patterns of households and businesses. One key feature of highly automated vehicles is enabling multitasking during vehicle operation, which would make the time lost in driving more productive. For example, travelers can do their work while being driven to their
office. Thus, the value of individuals’ travel time might change as travelers may not mind spending more time moving in a vehicle. In addition, having a robotic “Chauffeur” to assist in daily chores can reprioritize activities in the household. For instance, highly automated vehicles could pick up kids from school or groceries from the store.

Other characteristics of the new technology, such as the potentially higher safety and lower costs of SAVs, could affect other high-level decisions such as owning a vehicle (34-38). Households may require fewer owned vehicles since those vehicles can drive themselves and efficiently serve multiple members of the households. The new shared-automated service may eliminate the need to own a vehicle altogether if the service proves to be reliable and affordable. Vehicle ownership of businesses can also be affected by automated vehicle technology. With safer, more efficient, and more sustainable distribution, businesses may require fewer vehicles to deliver goods to their clients. They might also share automated vehicles for delivery to achieve higher utilization and lower costs.

**Travel Choices**

On the tactical level, some of the new features of AVs - mainly the capability to multitask during the trip - can affect individual trip decisions (16) such as mode choice, route choice, and departure time. The usual assumption is that human drivers choose travel routes that minimize their travel time according to the best information they have available. However, travel times are dynamic, depending on prevailing traffic conditions, and may not be readily available to non-connected drivers at the time of choosing their routes. Connectivity can affect route choice in several ways. One way would be for connected vehicles to act as probes to traffic conditions and
share that information with other connected vehicles. This would enable more accurate estimates of travel times and shortest routes. Another way would be for automated or connected vehicles to reroute themselves while moving towards a destination based on developing traffic conditions, which can lower costs and travel time.

As for mode choice, connectivity allows for new mobility tools and better intermodal integration. Travelers would be able to use multiple modes conveniently, for example, using public transit and a shared-automated-vehicle (SAV) service. Users may also shift to entirely different modes, like solely using SAV instead of driving personal vehicles or transit.

Departure times can also be affected by CAV technologies. With less variable travel times, travelers may not need to leave much earlier than they should to account for unforeseen delays. Additionally, travelers living in the same household can share an automated vehicle and coordinate their departure times. For example, parents can send their children to school in the highly automated car while getting ready to leave for work before the car comes back

2.1.2 Supply Changes

The major supply changes expected from the deployment of CAV systems are 1) the emergence of new mobility options and 2) changes to the infrastructure to enable wireless telecommunications. The new mobility services will mainly be in the form of a Shared Automated Vehicle (SAV) and/or hybrid systems enabled by SAVs (36; 39; 40). It also includes automated truck systems with potentially disruptive impacts to the trucking industry (41). The aforementioned
changes are captured in the supply component of the CAV AMS framework as illustrated in FIGURE 3.

FIGURE 3 The supply changes component of the general CAV AMS framework (1)

New mobility options

The rapid development in wireless telecommunication technologies and the high adoption rate of those technologies have enabled radically new forms of mobility and opportunities for multi-mode integrations that were not possible or thought of less than 20 years ago. Most AMS tools, for example, failed to predict current ride-hailing services, such as Uber and Lyft, which
were only enabled by advancements in positioning, telecommunication, and handheld computing technologies.

The current development in the areas of artificial intelligence, robotics, CAV systems, and the internet of things will probably cause even more radical changes to the forms of mobility that travelers are used to. Aside from the extremely futuristic modes such as flying cars or the Hyperloop, the most anticipated mode enabled by the aforementioned technologies is SAV fleets.

SAV fleets or their hybrid systems will play a key role in expanding mobility as a service (Maas) and create integrated forms of mobility in the future. For example, a new mode can be an integrated transit-SAV system where the latter serves as a first/last mile connection. While shared vehicle fleets are not an entirely new form of mobility, transport network companies (TNC) such as Uber and Lyft already offer this service, SAVs has two main differentiating features: (1) the automated driving behavior of vehicles is different from the human driving behavior, and will likely impact the overall performance of the system, and (2) the mobility service owner would have full control over the system, unlike services using human drivers, and can optimize the service to serve different objectives such as minimizing costs or maximizing quality. These two features have the potential to increase the SAV market share and competitive advantage against other modes.

Furthermore, connectivity can enable better integration of multiple modes for improved mobility. One specific case is public-private partnerships to solve the first/last mile problem of access to transit systems. Through a connected platform, for example, TNCs can integrate their services with transit systems to provide better accessibility and more convenient transfers.
Consequently, improving transit services through such partnerships can increase ridership and potentially reduce the need to use private cars.

In addition to personal mobility, the logistics industry could be one of the early adopters of automated vehicle technology as it promises improved safety, sustainability, and efficiency of goods movement \(^{(41)}\). Using automated trucks improves safety by reducing human errors. As for sustainability, automated truck technology can lower emissions by potentially improved fuel consumption. Finally, in the long term highly automated trucks can increase operational efficiency as machines, unlike drivers, do not need breaks between trips.

*Infrastructure changes*

The second major supply impact as a result of CAV technology deployment is the potential changes to the infrastructure to enable wireless telecommunications. This unique feature of CAV systems is often missing in existing CAV AMS capabilities. Reliable wireless telecommunication is not only essential for the operation of CAV technologies but can also affect the driving behavior of CVs. Most AMS tools, especially SAV fleet modeling ones, assume that all vehicles are connected and the central dispatcher has full information regarding the location of all vehicles, requests, origins, and destinations. This may not be the actual case in practice.

Furthermore, V2I technology, depending on the type of technology, is likely to be deployed in strategic locations due to its high costs. This will impact the operations of SAV fleets that rely on a central dispatcher to assign vehicles. Furthermore, wireless telecommunications, even the most advanced technologies to date, may not reliable at all times. It may suffer from outages,
disconnections, or poor signals, especially at severe weather conditions. Similar reliability issues involve the positioning of vehicles such as lost GPS signals inside tunnels.

For the abovementioned reasons, having an abstract representation of wireless telecommunications in CAV AMS systems is important for a realistic representation of new mobility options and evaluating the telecommunication impacts on driving behavior. In this methodological framework, wireless telecommunication technologies (V2I/V2V/V2X) are integrated within the network representation. For DSRC telecommunication technology, the representation would include communication ranges which affect the information flow between connected agents (travelers, vehicles, infrastructure).

2.1.3 Operational Performance

The most direct impact of CAVs on network performance will result from the operational performance characteristics of the vehicles in the traffic stream, and the control algorithms enabled by and deployed with varying degrees of V2V and V2I connectivity (42). CAV systems are expected to improve different performance aspects (43) of transportation systems including safety (44; 45), mobility (46; 47), and sustainability (48). The technology promises to reduce accidents that are caused by humans, improve road capacities by driving safely at higher densities (49), and improve traffic control whether on freeways (17; 19; 22; 23; 30; 33; 50) or intersections using advanced wireless telecommunication technologies (51-63). While greatly dependent on decisions made in the commercial marketplace, public agencies, and regulatory bodies, understanding and modeling these impacts under a given set of assumptions about technological features, deployment
scenarios, and control measures is an essential AMS requirement that lies mostly in the realm of traffic physics.

To fully capture the traffic impacts of CAV systems, AMS models should capture the heterogeneous interactions between different driving behaviors. First, there will be isolated manual drivers who have relatively higher reaction times and risks of driving errors. Second, there will be connected and well-informed drivers who are more aware of their surroundings and presumably with better reactive behavior. Finally, there will be the new driving behavior with the introduction of highly automated vehicle, which can also be connected through wireless telecommunications. This behavior would heavily depend on the equipped sensors and the control algorithms installed by car manufacturers in addition to the additional information that can be received through connectivity.

The operational performance component of the envisioned CAV AMS system, in FIGURE 1, is an integrated traffic-telecommunication simulation platform that can simulate mixed traffic conditions under different operational assumptions and scenarios. The performance component, illustrated in FIGURE 4, includes four types of driving behaviors: (1) isolated-manual, (2) connected-manual, (3) isolated-automated, and (4) connected-automated. It also includes a wireless telecommunication component which specifies the performance of the communication systems that is relevant to transportation system performance. Finally, the tool includes a component to simulate the heterogeneous interactions among the different driving behaviors depending on the assumed connectivity/automation levels and the implemented control algorithms.
As inputs to the integrated simulation platform, the framework includes demand patterns and the system configuration, which are outputs of the strategic-level analysis as discussed in the general framework (FIGURE 1). In addition, external factors (e.g. weather), logic for controlling automated vehicles, and the agency’s communication protocols are considered in the integrated simulation platform. Finally, the simulation tool outputs pre-defined, as well as user-defined, performance measures to evaluate the impacts of CAV technology on the system’s performance. The remainder of this section discusses the major components of the operational performance component in further detail.
Wireless Telecommunication and Sensors

The CAV AMS modeling system needs to provide for an appropriate level of representation of the effects that the enabling technologies for CAVs will have on the behaviors of the vehicles and their interactions with the transportation management functions. The most important of these technologies are the telecommunications and environment perception (sensing) technologies, but these are also closely coupled with positioning technologies. These technologies typically function on time scales much shorter than the time scales associated with vehicle motions, but that does not mean that they need to be modeled at those very short time scales (which would have adverse consequences for computational efficiency). The phenomena that influence their
performance are often very different from the phenomena that are represented in transportation network models (for example, ambient lighting conditions and atmospheric conditions that affect radio wave propagation and visibility, including disturbances such as electrical storms and sunspots). These considerations point toward the need for simplified models focused on the aspects of performance of sensors and communication systems that directly influence vehicle performance.

*Connected Manual Driver Behavior*

Connectivity extends drivers’ perception of their surrounding environment beyond the visual scanning of isolated drivers, leading to a more responsive driving behavior (4). Depending on the type of communication, V2V and V2I provide different information to drivers and affect their behavior accordingly. V2V provides information on vehicle movement and location, such as speed and acceleration of downstream vehicles, which increases drivers’ awareness of downstream traffic conditions and improves their responsiveness (lower reaction time). V2I, on the other hand, provides information on road conditions, weather, and TMC decisions (e.g. express lanes) influencing the drivers’ strategic decisions such as route choice and departure time.

Because of the above-discussed influences of connectivity on driving behavior, the proposed operational performance component explicitly distinguishes between the two manual driving behaviors (connected vs. isolated) and uses different acceleration/lane changing formulations to model them.

*Automated Driving Behavior*
A great diversity of driving automation systems needs to be represented by the CAV performance simulation tools since the systems under development and consideration vary widely from each other. The main dimensions to characterize driving automation systems are:

1. SAE levels of automation, defining which roles are performed by the automation system and which roles are performed by humans
2. Degree of coordination or cooperation – is the system autonomous or does it rely on V2V or I2V or more general V2X information?
3. Operational design domain (ODD) – the specific conditions under which the driving automation system is designed to function, including roadway type, traffic conditions and speed, geographic locations (boundaries), weather and lighting conditions, condition of pavement markings and signage, availability of other necessary supporting infrastructure features, etc.

The levels of automation are defined precisely in the SAE J3016 Recommended Practice document (64). This is an important reference that all modelers and designers of automated systems should study. The simplified version of the SAE J3016 classification criteria can be distilled into the answers to the following questions:

1. Does the driving automation system perform either the longitudinal or the lateral vehicle motion control task in a sustained fashion, but not the other? If yes, it is a Level 1 system. Many of these are already available to the public (such as adaptive cruise control or lane tracking systems).
2. Does the driving automation system perform both the longitudinal and lateral vehicle motion control tasks in a sustained fashion simultaneously? If yes, it is at least a Level 2 system. Some Level 2 systems are already available on premium vehicles and many more are under development, but at this level they still require the driver to continuously monitor the system performance and the driving environment for potential hazards.

3. Does the driving automation system also perform object and event detection and response? If yes, it is at least a Level 3 system. At this level, drivers can temporarily divert their attention away from the dynamic driving task to perform other tasks (such as reading or web surfing), but they need to be available to resume driving when the system requests help. None of these systems have been brought to the market yet, and there is controversy within the industry about whether it is possible to make such a system safe.

4. Does the driving automation system also perform the dynamic driving task fallback function, ensuring recovery from all internal faults or external hazards without requiring driver intervention? If yes, it is at least a Level 4 system. Some of these systems may not require any driver if their operation is confined to locations where the Level 4 operations can be guaranteed to function all the time (such as airport people movers in physically protected rights of way). A wide range of Level 4 systems are under development, but the critical aspect that needs to be defined clearly is the unique operational design domain for each system.
5. Is the driving automation system limited to operations within a specific operational design domain (ODD)? If it is not limited by an ODD, but is capable of driving safely under the full range of conditions in which humans can drive safely, it is a Level 5 system. That is a very long-range prospect, not something that needs to be planned for within the foreseeable future.

The level of automation of any specific driving automation system determines which aspects of its behavior need to be modeled as automated and which aspects need to be modeled as normal human driving behavior (using baseline driver car following or lane changing models).

*Heterogeneous Traffic Interactions*

The introduction of CAV on the road will lead to different types of drivers sharing the same transportation facility. To evaluate the flow impact of the traffic interactions among the heterogeneous driving styles and since the actual CAV market share is a variable that can be chosen to have many different values, the operational performance component should be able to simulate multiple scenarios at different CAV market penetrations. This would help planners and policy makers prepare for the impacts of CAV in the short and long-term as the market penetration of CAV is expected to start small and grow as the technology matures.

In addition to evaluating mixed traffic, the operational performance component can be used to evaluate special control algorithms designed for the CAV environment. For example, the simulation tool can be used to evaluate a special speed harmonization algorithm which utilizes more accurate traffic volumes available through V2V communications, and sends the speed limit
information directly to connected drivers instead of using fixed signs. Another example of special algorithms is a queue warning system which use vehicle locations through V2V communications to accurately detect queues, and directly warns drivers upstream of queues instead of using fixed signs.

**Network Integration**

Evaluating the network-wide impacts of CAV systems requires capturing the interactions of different agents in a network context. Those agents include CAVs, travelers, mobility service providers, transit and network managers, freight shippers and carriers. To capture these interactions, model *platforms* are required, integrating various components relevant to the questions being asked. Platforms in this context are primarily conceptual analytical constructs that are embedded in a software tool. They typically entail a collection of models representing interacting agents or processes. In this case, the CAV AMS system would be a platform that integrates a collection of supply, demand, and performance models to represent the behavior of CAV systems and their impacts on transportation systems, as illustrated in FIGURE 5. Platforms also typically offer a foundation upon which additional capabilities may be built, albeit with varying degrees of difficulty and effort.
FIGURE 5 The network integration component of the general CAV AMS framework (I)
2.2 Evaluating the Operational Impacts of Connected and Automated Vehicle Systems

This section provides a review on some of the automated driving behavior models used for traffic simulation in addition to some of the prior/ongoing work on CAV-related traffic control and policies, including speed harmonization, reserved lanes for automated vehicles, and queue warning. Early automated driving models focus on automating car-following as an assistant system to the driver. To do so, the models assume that vehicle uses basic sensors to get information about relative speed and distance to the leading vehicle. Using such information, the vehicle is able to adjust its speed automatically keeping a safe distance from the leading vehicle. Such systems are called advanced cruise control or adaptive cruise control. Other models extended advanced cruise control models to use V2V and V2I communication technology to predict the traffic state ahead of the vehicle and create platoons that can travel at closer relative distances. Such models are called cooperative adaptive cruise control. TABLE 1 provides a summary of selected studies on the impacts of automated cruise control systems on traffic flow. The studies are described further in the following subsections.
### TABLE 1 Summary of selected studies on the impacts of Automated Intelligent Cruise Control (AICC)/Adaptive Cruise Control (ACC) and Cooperative Adaptive Cruise Control (CACC) on traffic flow

<table>
<thead>
<tr>
<th>Study</th>
<th>Model</th>
<th>Connectivity</th>
<th>Major Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ioannou and Chien (33)</td>
<td>AICC</td>
<td>No</td>
<td>AICC can lead to smoother traffic flows and larger traffic flow rates, and can outperform human driving in emergency cases</td>
</tr>
<tr>
<td>Van Arem et al (17)</td>
<td>AICC</td>
<td>No</td>
<td>AICC can reduce the number of shockwaves generated in traffic stream and the number of vehicles inside them In high demand scenarios, AICC can deteriorate flow rate</td>
</tr>
<tr>
<td>James et al (22)</td>
<td>ACC; different models</td>
<td>No</td>
<td>ACC has a minor impact on traffic flow at low market penetrations while it has a negative impact at higher market penetrations</td>
</tr>
<tr>
<td>Van Arem et al (23)</td>
<td>CACC</td>
<td>Yes</td>
<td>Traffic flow improves at high demand and CACC market penetrations while it deteriorates at low market penetrations</td>
</tr>
<tr>
<td>Vander Werf et al (19)</td>
<td>ACC/CACC</td>
<td>No/Yes</td>
<td>ACC systems have minimal effect on highway capacity even at high market penetrations while CACC systems have a significant effect on highway capacity that proportional to the market penetration of the technology</td>
</tr>
<tr>
<td>Shladover et al (30)</td>
<td>ACC/CACC</td>
<td>No/Yes</td>
<td>ACC is unlikely to produce a significant increase in capacity</td>
</tr>
<tr>
<td>Melson et al (50)</td>
<td>CACC – Network Dynamic Traffic Assignment</td>
<td>Yes</td>
<td>Travel time reductions proportional to demand levels and significant reduction in congestion due to CACC</td>
</tr>
</tbody>
</table>

### 2.2.1 Automated Intelligent Cruise Control/Adaptive Cruise Control

In one of the early work on automated driving models, Ioannou and Chien (33) developed an Automated Intelligent Cruise Control, also referred to as Adaptive Cruise Control, system for automatic car-following where they examined the system’s effect on traffic flow and compared its performance with human driver models. The AICC system does not exchange information with other vehicles, but has access to relative speed and velocity with respect to the leading vehicle. To
eliminate oscillation effects, the authors used a safe distance separation tool that is proportional to the vehicle velocity (constant time headway) and designed the system accordingly. The constant headway was calculated using a worst-case stopping scenario.

The authors used simulation experiments to compare AICC with three human driver models. The oscillations and long settling times observed with human driver models are non-existent in automatic vehicle following. Results indicated that automatic car following can lead to smoother traffic flows and larger traffic flow rates due to automated vehicles driving with shorter safety spacing, and less reaction times. The authors also concluded that AICC could outperform human driving models in different emergency cases like emergency stopping and cut-ins. More information on the control logic and simulation experiments are found in the paper.

In another work on modeling Automated Intelligent Cruise Control, Van Arem et al (17) proposed an AICC system that automatically maintains a desired speed of the vehicle taking into account a minimal headway with respect to the leading vehicle. As in the case of Ioannou and Chien’s work (33), the system is assumed to be independent and disconnected from other vehicles or road-side systems. Furthermore, the driver is assumed to take over control of in case of emergencies. The authors used the simulation model MIXIC to study the potential impact of AICC on traffic. The model assumes that relative speed and distance is obtained from a basic sensor.

Simulation results showed that AICC can reduce the number of shockwaves generated in traffic stream and reduce the number of vehicles inside those shockwaves, indicating a more stable traffic flow. However, in some simulated scenarios where traffic demand was high, results showed
that AICC can deteriorate traffic performance. On the other hand, a low AICC penetration had no significant effect on traffic.

James et al (22) assessed the impacts of ACC on traffic flow using four ACC car following models programmed into the simulation platform VISSIM. The ACC models tested were MIXIC (28), IIDM (49), Path empirical (32), and Delft empirical (20). Furthermore, the models were calibrated using data collected by vehicle with a 2013 Cadillac SRX with a production ACC-enabled system while following a human-driven 2013 Cadillac SRX in northern Virginia. The results show that models tested are different in their sensitivity to calibrated coefficients. In addition, the simulations show that ACC has a minor impact on traffic flow at low market penetrations while it has a negative impact at higher market penetrations emphasizing the importance of connectivity in automated cruise systems.

2.2.2 Cooperative Adaptive Cruise Control

Van Arem et al (23) extended the concept of AICC to include V2V communications so that automated vehicles can follow leading vehicles at a closer distance. In addition to knowing relative distance and speed, V2V communications allows vehicles to coordinate speed changes, exchange precise speed information, accelerations, warning of forward and hazards, and maximum braking capabilities. The authors used the traffic simulation tool MIXIC to study the Cooperative Adaptive Cruise Control (CACC) effect on traffic characteristics.

Simulation results showed an improvement of traffic-flow stability and a slight increase in traffic-flow efficiency. The traffic flow especially improves in conditions with high-traffic volume
and when high fractions of the vehicle fleet are CACC equipped. At low-CACC presence (< 40%), results indicated a degradation of performance demonstrated by lower speeds, higher speed variances, and more shock waves. The system has a negative effect on traffic safety in the merging process; close CACC platoons prevent other vehicles from cutting in resulting in an increasing number of removed vehicles due to conflicts. As for shockwave, simulations showed a decrease in the number of shockwaves before a lane drop when high number of CACC-equipped vehicles are present.

In another work by Vander Werf et al (19), the authors studied the effects of Adaptive Cruise Control (ACC) and CACC on highway traffic flow capacity using a Monte Carlo simulation approach. Three types of vehicles were simulated in the study: 1) vehicles driven by humans, 2) vehicles equipped with ACC system to control speed with 1.4 s time gap, and 3) vehicles equipped with CACC system enabled by V2V and using a time gap of 0.5 seconds. Furthermore, the two automated cruise systems were simulated for different scenarios by varying market penetrations.

The study’s results show that ACC systems have minimal effect on highway capacity even at high market penetrations (7% increase in capacity at most.) On the other hand, CACC has a significant effect on highway capacity that is proportional to the market penetration of the technology. At full CACC market penetration, for example, the highway capacity can increase to more than double the capacity of the base case (without ACC or CACC systems)

Shladover et al (30), also studied the effect of Adaptive Cruise Control - ACC and CACC on highway capacity using the micros-simulation tool AIMSUN. The authors used the distribution of time gap settings by drivers that participated in a real field experiments previous to their study.
The authors simulated four types of vehicles: manual vehicle with driving behavior represented by the NGSIM oversaturated flow model, ACC vehicle with driving behavior represented by a simple first-order control model, Here-I-am (HIM) vehicle which constantly broadcasts its location, CACC vehicle that uses its capability if it follows HIA or CACC vehicle and acts as a normal ACC vehicle otherwise.

Results showed that ACC is unlikely to produce a significant increase in capacity as drivers are comfortable with driving gaps that are similar to the gaps drivers choose when driving manually. CACC, however, showed a potential for significant increase in capacity at high market penetration. This is due to drivers being more confident to follow vehicles shorter gaps due to higher dynamic response of CACC over ACC.

In addition to the abovementioned microsimulation approaches, Melson et al (50) studied the effect of CACC at the network level by incorporating CACC into the link transmission model (LTM) for dynamic network loading. As a first step, the authors derived the CACC flow-density relationship (fundamental diagram) of CACC from the MIXIC car following model. After verifying the fundamental diagram with the observed speeds and flows using the simulation platform VISSIM, the authors created a network loading model using the aforementioned fundamental diagram in LTM.

Comparing DTA and MIXIC microsimulation on a subnetwork, both models predicted travel time reductions (up to 32%) with increasing demand as a result of CACC. The authors also tested CACC on two larger networks: a 28-mile corridor of I-35 near Austin, Texas where all vehicles were assumed to be equipped with CACC, and the Round Rock Network where one
CACC lane was added. Results of both networks shows a significant reduction in congestion due to CACC, however, the Round Rock network results indicate an increase in the overall travel time due to rerouting. This implies the importance of including user route choice in the DTA analysis of CACC.

2.2.3 Speed Harmonization in a Connected Vehicle Environment

Speed harmonization is a traffic control strategy that adjusts the speed limit of a freeway section based on prevailing traffic conditions (61; 62; 65). The strategy helps mitigate shockwave formation, damping its propagation, and minimize the spatial variance of speed to homogenize traffic and accelerate the recovery from a traffic breakdown (2; 61; 65-67). Two critical elements of this strategy are 1) shockwave detection and 2) speed limit broadcasting to upstream vehicles. Traditional speed harmonization strategies rely on sensors embedded in the infrastructure to detect shockwave formation and on fixed-in-location dynamic signs to update suggested speed limits.

Connected vehicle technology provides new opportunities for shockwave detection and broadcasting speed limits. CAVs can act as probe vehicles that monitor their surrounding traffic conditions and broadcast that information to central traffic management systems. Those systems use the new information received to identify shockwave locations more accurately and directly send new speed limits to connected vehicles upstream of shockwave locations. This improved speed harmonization strategy allows a greater range of effectiveness and gradation in displayed speed limits than conventional strategies (2; 61).
Automated vehicle technology, on the other hand, can also help dampening shockwaves through controlling velocity of automated vehicles in the traffic flow (25). In a field experiment conducted by Stern et al (25) on a closed ring road, the authors found that one automated vehicle can control the flow of at least 20 human controlled vehicles around it. The speed-controlled automated vehicle can substantially reduce the speed variation among vehicles, excessive braking, and fuel consumption.

In Ma et al’s review (65; 68) of recent speed harmonization algorithm developments, summarized in TABLE 2, the authors categorized the CAV-enabled algorithms into: 1) algorithms that use shared information with CV system (60-63) and 2) algorithms that control vehicles equipped with CAV systems (26; 69). The results of the studies reviewed shows the effectiveness of CAV-enabled speed harmonization in delaying/ or dampening traffic oscillations in addition to improving safety and sustainability.
TABLE 2 Summary of recent studies on speed harmonization applications enabled by CAV systems - Ma et al (65; 68)

<table>
<thead>
<tr>
<th>Study</th>
<th>Comm.</th>
<th>Input</th>
<th>Control Algorithm</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lu et al (60)</td>
<td>V2I</td>
<td>Segment speeds, detailed trajectory-level data not necessary</td>
<td>Reduce speed limits of freeway segments upstream of a bottleneck in proportion to the observed bottleneck speed if vehicle flow throughputs are above the bottleneck capacity</td>
<td>Works for a corridor and a freeway network with multiple bottlenecks</td>
</tr>
<tr>
<td>Talebpour et al (61)</td>
<td>V2I</td>
<td>Detailed microscopic vehicle trajectory</td>
<td>A wavelet-transform based algorithm to detect formation of perturbations; a cognitive risk-based microscopic simulation model was adopted to account for human behavior; a reactive speed limit was selected to implement SH reactive speed limit was selected to implement SH</td>
<td>Effectively delay or eliminate traffic breakdown and improve traffic safety even at a low penetration rate of 10%</td>
</tr>
<tr>
<td>INFLO project (62)</td>
<td>V2V/V2I</td>
<td>Speed measured from connected vehicles and infrastructure-based sensors</td>
<td>Group freeway sub-links with similar recommended speeds to produce harmonized speeds</td>
<td>SH effectiveness depends upon driver compliance</td>
</tr>
<tr>
<td>Li et al (66)</td>
<td>V2V</td>
<td>Leading vehicle’s input</td>
<td>CAV car-following rule</td>
<td>Effectively suppress development of oscillation and consequently mitigate fuel consumption and emission</td>
</tr>
<tr>
<td>Wang et al (69)</td>
<td>V2I</td>
<td>Aggregated traffic state</td>
<td>Use aggregated traffic state information to detect formation of congestion at a bottleneck; each CAV processes the VSL signals from the central control unit individually</td>
<td>The connected VSL and vehicle control system improves traffic efficiency and sustainability, i.e., total time spent in the network and average fuel consumption rate are reduced</td>
</tr>
<tr>
<td>Yang and Jin (63)</td>
<td>V2I</td>
<td>Individual vehicle’s information</td>
<td>Advisory speed limit is calculated by each individual vehicle and then averaged among green driving vehicles</td>
<td>When 5% of the vehicles implement the green driving strategy and the communication delay is 60 s, the fuel consumption can be reduced by up to 15%</td>
</tr>
<tr>
<td>Ahn et al (70)</td>
<td>Radar and V2V</td>
<td>Topographic information, the spacing between the subject and lead vehicle, and a desired (or target) vehicle speed and distance headway</td>
<td>Use a rolling horizon-based optimization approach to control vehicle speed within a preset speed window in a fuel-saving manner</td>
<td>Simulated fuel savings in the range of 27% are achieved with an average vehicle spacing of 47 m along a study section of Interstate 81</td>
</tr>
</tbody>
</table>
2.2.4 Dedicated Lanes for Automated Vehicles

One way to minimize the interactions between automated vehicles and regular vehicles in mixed traffic conditions is dedicate a number of lanes for the former. Similar to HOV lanes, this setup can attain higher traffic throughputs since automated vehicles would be able to drive at higher densities than regular vehicles (lower safe distances). The expensive option of this strategy is to add a new lane to a freeway that is dedicated to automated vehicles. A less expensive and more readily implemented option is to reserve one of existing lanes to AVs. This approach, however, could risk higher congestion in regular lanes.

To explore the potential implications of dedicated AV lanes, Talebpour, Mahmassani and Elfar (71) used a microscopic simulation platform to test an operational scenario of a 3.5-mile section of a four-lane freeway in the Chicago region. Three distinct operational policies were tested in conjunction with reserving the leftmost lane for automated vehicles: (1) mandatory use of the reserved lane by autonomous vehicles, (2) optional use of the reserved lane by autonomous vehicles, and (3) limiting the autonomous vehicles to operate autonomously in the reserved lane.

The findings of these investigations suggest that the optional use of the reserved lane without any limitation on the type of operation can improve congestion and reduce scatter in the fundamental diagram. In contrast, limiting autonomous vehicles to the reserved lane and preventing autonomous operation in regular lanes could significantly increase congestion and result in breakdown formation. In particular, mandatory lane-changing maneuvers of autonomous vehicles are the main source of shockwave formation. The analysis is extended to higher overall flow levels (beyond those currently observed on that facility) to explore the minimum and optimal
threshold levels for introducing such reserved lanes. In this case, reserving one of the four lanes for autonomous vehicles is only beneficial at market shares above 30%. Furthermore, travel time reliability analysis revealed that optional use of the reserved lane can yield the most benefit.

In a working paper by Su et al (72), the authors investigated introducing CACC vehicles using three lane management strategies. The first strategy is High-Occupancy-Vehicle (HOV) only where the left-most lane accepts only HOV vehicles. In the second strategy, the left-most lane is open to both HOV and CACC vehicles. In the third one, the managed lane is only open to CACC vehicles. The three scenarios were simulated using a microscopic freeway simulation platform VISSIM and the MIXIC model for dynamic CACC operations in a 14-mile section of Interstate-66 near Washington, DC.

The simulation results showed that the dedicated CACC lane’s capacity can reach as high as 3800 vphl because of having higher and more stable speeds than general purpose lanes. However, for low market penetrations, the analysis shows that dedicated lanes are inefficient. For market penetration (MPR) below 25%, HOV + CACC is the best strategy where capacity increases by 10% for 25% MPR over the base case. At 45% MPR, the CACC only strategy was the best out of the three where capacity increased by 10% for the whole corridor.

2.2.5 Queue Warning in a Connected Vehicle Environment

Queue warning, in its original form, has been used in multiple locations around the world. It provides drivers sufficient warning regarding downstream queues in order to reduce speed safely, change lanes, or use alternative routes (73). These smooth changes in time could minimize
or eliminate rear-end collisions. However, connectivity addresses the current system’s limitations to further improve its effectiveness and efficiency. V2I communications can provide information on current location of vehicles. Therefore, queues can be located more accurately and in real time compared to relying on fixed detectors. It also allows drivers to receive warnings faster and more reliably. Connectivity also helps drivers receive those warning in the right place at the right time, compared to relying on fixed message boards that might be too close or too far depending on the queue location. It may also help predict queues in advance based on real time data provided by downstream vehicles and weather conditions. FIGURE 6 shows the queue warning concept presented by Mahmassani et al. (66) in their report on intelligent network flow optimization.

FIGURE 6 Queue Warning Concept

The literature on modeling queue warning and investigating the effectiveness of this application is limited. Pesti et al. (74) used macroscopic traffic simulation to evaluate different design alternatives of a dynamic queue warning system at a freeway work zone with lane closure. Their objective was to assess the expected performance of a dynamic queue warning system deployed at a freeway, evaluate the sensitivity of the system to key design parameters, and
determine the most effective combination of these parameters. The authors found that queue warning with less spacing between detectors can be more accurate at detecting queue-ends. The authors recommend a 35mph queue detection speed. Finally, they concluded with a recommended set of design parameters for an effective queue warning system. In another study, the Washington State Department of Transportation (75) conducted a feasibility study of different active traffic management techniques including speed harmonization and queue warning. In their study, they estimated the capital cost of implementing each technique per location and used a microscopic simulation model to evaluate the benefits. Their simulation results suggest that implementing queue warning could save travelling time by reducing delay. Khan (76) presented a queue-end warning system that predicts queue ends and notifies drivers of the predicted queue-end location using portable variable message signs. The system is a combination of traffic sensors and a queue-end prediction algorithm based on an artificial neural network model.

Traditional queue warning systems are being used in different cities around the world. Pesti et al. (77) evaluated a traditional queue warning system deployed by the Texas Department of Transportation on US-59 and IH-610 near Houston, TX. The system used video cameras to determine vehicles speed, static message board, and two flashing beacons that are activated based on congested conditions. Their before/after evaluation criteria were based on accident history, traffic conflicts, and vehicles speed. The authors found that an insignificant increase was found after deploying the system. However, they also mention that their evaluation period is much less than required for a safety evaluation. As for vehicle conflicts, they found that the conflicts have decreased after deploying the system, although not significantly. Finally, they found that vehicles speed slightly decreased on IH-610 but slightly increased on US-59. The authors recommended
using lane selection messages, adding another speed detection camera on IH-610 for more accurate prediction, and informing drivers of actual speeds downstream using advisory speed messages. Wiles et al. (78) examined current queue waning practices inside the U.S. and internationally. One of the queue warning systems presented in their paper is an automated system in Oslo, Norway that utilizes video cameras to detect vehicle speeds while variable message system is used to alert drivers. The authors concluded that an effective queue warning system should adjust to fluctuating queue conditions.

### 2.3 Chapter Summary

This chapter provides an overview of the potential strategic and operational impacts of CAV systems. It also provides a detailed review of previous work on modeling the operational impacts of CAV systems; the focus of the new CAV control strategies developed in the dissertation. The chapter reviewed efforts at modeling CAV new driving behavior such as ACC and CACC systems. It also reviewed efforts at developing traffic control applications that utilize the new technology such as speed harmonization and queue warning. Finally, the review also examined potential policies such as assigning dedicated lanes to CAVs.

The lack of actual data on the behavior of the new systems is the major limitation of most of the studies reviewed. Therefore, the majority relied on best-guess assumptions regarding the new systems in their choice of model parameters. A limited number of studies, however, used data generated from field experiments to calibrate AV models. The findings of the reviewed studies suggest that CAV systems have a positive impact on the stability and rate of traffic flow that is
proportional to CAV market penetrations. Furthermore, the new technology can potentially improve traffic control algorithms such as speed harmonization.
3. FRAMEWORK FOR PREDICTIVE TRAFFIC MANAGEMENT AND CONTROL STRATEGIES OF CONNECTED AND AUTOMATED VEHICLES

The potential changes brought by CAV technology impact traffic control applications in multiple ways. The technology introduces new driving behaviors that are expected to improve the stability and performance of traffic flow \(^2\cdot 4\). Those behaviors will create unprecedented interactions between regular human drivers and CAVs that need to be captured by future control strategies. CAVs would also generate detailed vehicle trajectories that enable the estimation of new traffic properties, such as the speed variation of vehicles \((79)\), which are typically difficult to estimate using traditional road sensors. Estimating those properties allows tracking of finer traffic patterns that can improve the prediction of traffic states \((79)\) and therefore the effectiveness of traffic control. The wireless telecommunication capability of CAVs offers more flexibility in executing various control strategies. Using V2V or V2I communications, control actions (e.g. updated speed limit) can be directly sent to CAVs without the need for fixed road signs. It would also improve the range over which control is applied since it is not constrained by road sensor locations.

This chapter provides a framework for developing predictive traffic control strategies that utilizes the new capabilities and generated data of CAVs. The framework, illustrated in FIGURE 7, serves as a roadmap for the development of the CAV predictive strategies in this dissertation. The framework consists of three main components: 1) traffic monitoring, 2) traffic state prediction, and 3) control strategy. The traffic monitoring component describes how the detailed vehicle
trajectories broadcasted by CAVs can be used to estimate traffic properties and track shockwaves (transitions in traffic state) without relying on road sensors. The traffic state prediction describes how traffic properties can be estimated through CAV data and be used for predicting the future traffic state, a key element in traffic control. Finally, the control strategy component describes two ways in which control actions (e.g. new speed limit, mandatory lane change) can be executed through CAVs: centralized and decentralized control. The three components are discussed in further detail below.
FIGURE 7 Framework for developing predictive traffic control strategies of connected and automated vehicles
3.1 Traffic Monitoring

One of the most promising features of CAVs is their ability to broadcast detailed information about their movement to other vehicles through V2V communications or to infrastructure systems through V2I communications. The information they broadcast can include, among other attributes, speed, location, acceleration, and lane-change intent. These broadcasted attributes, call them vehicle trajectories, enables traffic management systems to track finer traffic dynamics and to apply more effective control strategies. Therefore, the proposed framework for CAV traffic management strategies considers utilizing the aforementioned feature for traffic monitoring as one of its main components. The detailed vehicle trajectories can be used in two key areas: 1) estimating traffic properties and 2) tracking traffic dynamics.

3.1.1 Estimating Traffic Properties

Current traffic monitoring systems rely on embedded road sensors to estimate aggregated traffic properties such as flow and speed. Common sensors, such as the inductive loop detector, count the number of vehicles that passes over it. Other sensors such as the double loop setup or radars can measure average speed of vehicles over a fixed time horizon. Installing and maintaining road sensors is often expensive and their measurements can be inaccurate depending on factors such as weather and road condition. Furthermore, the aggregated nature of the data collected by those sensors loses a lot of information that could be used for more effective control of traffic.

The wireless telecommunication feature of CAVs provides new opportunities to collect more detailed traffic information without the need for road sensors, potentially reducing their
installation and maintenance costs. For example, the speed variation of vehicles is an excellent indicator for traffic disruptions that can be estimated from CAV trajectories (79). The number of lane changes is also an indicator for traffic stability that can be estimated through CAV trajectories. Another advantage for using CAVs to estimate traffic properties is the ability to collect information anywhere on a road segment of interest since this approach is not constrained by the location of embedded road sensors. This offers more flexibility in applying traffic control strategies, which in turn could improve their performance.

3.1.2 Tracking Traffic Dynamics

The aggregated traffic properties collected by road sensors lack information that is critical to track fine traffic dynamics such as traffic shockwaves, a phenomenon that signals the transition between traffic states (e.g. between congested and uncongested). Currently, traffic shockwaves are detected by tracking changes in estimated traffic flow and density, both of which are prone to estimation errors. Furthermore, shockwaves can only be detected if a road sensor is available near the shockwave when it forms which limits the shockwave’s detection ability and accuracy. CAV technology provides more detailed information that can be used to track shockwaves anywhere on a road segment of interest (61). It broadcasts information at the individual vehicle level providing higher quality estimates of traffic properties to track shockwaves such as the speed variation of vehicles. Other traffic dynamics that can be tracked through CAV trajectory data include, for example, crashes and unexpected slowdowns.
3.2 Traffic State Prediction

The traffic state transition into the congested state is affected by a combination of three key factors: 1) high traffic loads, 2) capacity reduction, and 3) disturbances caused by individual vehicles. The first two factors lead to traffic loads that are higher than the capacity of a transportation facility, which leads to congestion. The third factor disrupts the stability of traffic and acts as a catalyst to traffic congestion. Measuring all three factors is essential to predict traffic state transitions, which is the first step for an effective traffic control strategy. High traffic loads and capacity reductions can be measured by road sensors since they rely on aggregated measures of traffic flow and density. On the other hand, measuring traffic disturbances by individual vehicles requires collecting data at the individual vehicle level, which as discussed before, is possible through the wireless telecommunication capabilities of CAVs.

Collecting such information for real-time traffic applications would generate large quantities of data that are beyond the processing capability of traditional traffic management systems. Therefore, this framework considers machine learning as a powerful tool for traffic state prediction using the broadcasted vehicle trajectories by CAVs. Machine learning techniques are capable of recognizing patterns in big data sets, which makes them suitable for CAV traffic management applications. To build a prediction model for CAV applications, two key areas needs to be evaluated: 1) machine learning technique and 2) prediction type, i.e. whether it is performed online or offline.
3.2.1 Machine Learning Techniques

Machine learning is emerging as a key mobility component in the internet of things world. Its ability to take advantage of large data sets of different sources makes it a versatile tool for traffic prediction applications. For example, machine learning can be used to predict traffic congestion \((80; 81)\) or forecast traffic flow \((82-87)\) from mobile GPS records. Multiple techniques can be classified under the umbrella of machine learning such as neural networks, decision-trees, and support-vector machines. Those techniques differ in aspects such as prediction accuracy, model training time, and data transformation. Therefore, selecting the right technique depends on the application itself. For instance, real-time online applications require fast training time of models since those would be updated in real time. Accident prediction, on the other hand, would require higher accuracy since it involves saving lives.

3.2.2 Prediction Model Type

This framework considers two types of prediction models: 1) offline and 2) online. Offline prediction models are trained (calibrated) using historical data and are updated whenever a major change occurs to the transportation facility (e.g. added lane). These models are usable when training data is time consuming or in situation where traffic state changes are historically recurrent. For example, offline models can be used to predict traffic congestion during typical work day during rush hours. Online models on the other hand are trained using historical data but are re-trained in real-time to capture unexpected changes to the transportation facility. For example, these models can be used to predict traffic congestion for a transportation facility during a major event (e.g. concert of a sports game) that is expected to disrupt traffic.
3.3 Control Strategy

Utilizing the estimated traffic properties and state prediction generated by previous components, this component applies a traffic control strategy through CAVs in order to improve the overall performance of the transportation system. The control may be advisory in nature where, for example, connected drivers receive suggested actions such as lowering their speed. It may also be mandatory, as in the case of automated vehicles that fully comply with specific actions such as lowering their speed or changing their lane. An effective control strategy aims to address one or a combination of the three main causes of traffic congestion: (5: 88):

1. High traffic loads
2. Bottlenecks
3. Disturbances caused by individual drivers

High traffic loads occur when traffic demand exceeds the sustainable throughput of a road section. A typical example is rush hour congestion. Capacity reductions or “bottlenecks” may be permanent, such as on-ramps and off-ramps, or temporary such as traffic accidents or slow-moving vehicles. As for traffic disturbances, those refer to temporary perturbations in the traffic flow, such as those caused by lane-change maneuvers, abrupt braking, or long-lasting overtaking maneuvers of trucks. All of those factors acting jointly or separately increase the likelihood of traffic congestion associated with the flow breakdown phenomenon (89: 90). To reduce the likelihood of flow breakdown, a control strategy can be applied through CAVs in two main ways: 1) centralized and 2) decentralized. For centralized control, a traffic management system analyzes the state of the traffic and applies a control logic to improve its performance. For the decentralized version,
each vehicle assesses the state of the traffic ahead of it and applies a control logic based on the information it receives from other vehicles. This control, however, may not be optimal to the overall system as individual vehicles would not have an overview of the whole system’s performance nor the computational power to optimize for it. While both types of strategies may have the same objective, such as mitigating congestion, the main difference between them is in way those strategies are evaluated and executed—either through a central system or through individual CAVs. The two strategy types are discussed further below.

3.3.1 Centralized Traffic Control (V2I)

Centralized strategies mainly utilize V2I wireless telecommunication technology where CAVs are connected to a central traffic management system. This system analyzes the information received from all CAVs in a facility and applies the traffic control logic accordingly, for example, instructing CAV vehicles to slow down to delay the onset of flow breakdown-induced congestion. The centralized control logic can be applied to minimize individual traffic disturbances or reduce inflow as potential strategies to minimize traffic congestion. As for the former, the system would:

1) evaluate the state of the transport facility through information received from CAVs and detectors then
2) predicts future states using machine learning algorithms
3) broadcast advisory messages to CAVs (speeds, lane changes) that minimizes disturbance (SSD)

As for reducing traffic inflow in case of high demand, a centralized traffic control system would:
1) monitor traffic properties (speed, SSD, lanes changes, acceleration) using CAVs and imbedded detectors in infrastructure

2) predict breakdown location based on traffic measurements using machine learning algorithms

3) reduce traffic speed upstream of breakdown location (choose optimal speed, advisory distance, duration) to lower incoming traffic and accelerate breakdown recovery

3.3.2 **Decentralized CAV Traffic Control (V2V)**

In the case of decentralized strategies, CAVs adjust their behavior, such as speed or acceleration, based on the information it receives from other CAVs (V2V communication). This type of control can be fleet-oriented where a group of vehicles are connected, and they adjust their behavior (speed, acceleration, etc.) based on the information they share with each other, for example, a truck platoon. It can also be oriented towards individual vehicles that adjust their behavior based on the information they receive from other connected vehicles but not necessarily acting as a group. This type of control can be effective for minimizing traffic flow disturbances or traffic inflow (when necessary). For reducing traffic disturbances, each CAV would:

1) receive information from a cluster or fleet of CAVs within a detection/connection range

2) use individualized or group-based machine learning algorithms to predict the future state of clusters (SSD, location, etc.)

3) adjust their longitudinal and lateral driving behavior to minimize disruption in a cluster or fleet of vehicles, i.e. self-homogenize (e.g. speed up slow vehicles or slow down speeding ones)
As for reducing traffic inflow when demand is higher than capacity, each CAV would:

1) share information about its state to other vehicles within a cluster/fleet
2) predict the future state of traffic ahead of it using vehicle-specific or fleet-specific machine learning model
3) adjust its speed (speed, acceleration, etc.) to temporarily reduce incoming traffic in order to resolve the bottleneck using a vehicle-specific or fleet-specific control logic

While decentralized control can be used to reduce traffic disturbances or inflow when necessary, it may result in disturbances created by partial slowdowns or different breakdown location predictions based on the information received by vehicles.

### 3.4 Chapter Summary

This chapter provides a framework for developing traffic control strategies that utilize the new capabilities of CAV systems. The framework consists of three main components: traffic monitoring, traffic state prediction, and control strategy. The traffic monitoring component describes how the detailed vehicle trajectories broadcasted by CAVs can be used to estimate traffic properties and track shockwaves (transitions in traffic state) without relying on road sensors. The traffic state prediction describes how traffic properties can be estimated through CAV data and be used for predicting the future traffic state, a key element in traffic control. Finally, the control strategy component describes two ways in which control actions (e.g., new speed limit, mandatory lane change) can be executed through CAVs: centralized and decentralized control. This
framework provides the foundation for the CAV traffic control strategies introduced in the following chapters.
4. TRAFFIC SHOCKWAVE DETECTION IN A CONNECTED ENVIRONMENT USING THE SPEED DISTRIBUTION OF INDIVIDUAL VEHICLES

Traffic shockwaves reflect a transition from the free-flow traffic state to the congested state (6; 88) and entail a number of negative impacts in terms of safety, performance, and emissions (91). They can create unsafe conditions as drivers might have to decelerate sharply when approaching a stream of slow-moving vehicles. Additionally, the frequent acceleration/deceleration behavior of vehicles in a propagating shockwave (traffic jam) increases fuel consumption and emissions. In terms of performance, shockwaves result in a significantly lower throughput than the nominal capacity of the freeway (91; 92).

The traditional approach for detecting shockwaves on a freeway is to track changes in speed and density over space and time based on fundamental diagrams (88; 91; 93; 94) and in accordance with the Lighthill-Whitham-Richards first-order continuum traffic theory (95; 96). The main limitation of this approach is that density along a freeway is difficult to measure in practice using traditional data collection methods such as loop detectors. It is rather estimated from other measures such as occupancy. Furthermore, locating the starting point of shockwaves may not be accurate as tracking changes in density over freeway segments depends on the number of installed loop detectors and the spacing between them (i.e., data resolution).

Wavelet transformation is a different approach that was proposed by Zheng et al. (97) to detect freeway bottlenecks and oscillations. The technique is a time-frequency decomposition
approach that extracts useful information from stationary time-series data, and it has been used in different intelligent transportation system applications (97). In their study, the authors used wavelet transformation to identify the time and location of a bottleneck activation by tracking the changes in average speeds recorded by loop detectors. They also proposed another application of wavelet transformation where they identified traffic oscillations (shockwaves) by tracking the acceleration/ deceleration cycles of individual vehicle trajectories. Assuming that such trajectories will be available through connected vehicles technology, Talebpour et al. (3; 61) adopted a similar wavelet based approach to identify shockwaves and to activate a speed harmonization system.

Connected vehicles technology provides various opportunities for enhancing the efficiency, productivity, sustainability, safety, and reliability of transportation systems (2; 3). One of the potential applications of the new technology is generating detailed vehicle trajectories through Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communications. Connected vehicles can send out basic safety messages that include real-time updates of a vehicle’s status, including its location, speed, and acceleration to other vehicles or to the infrastructure and receive the same types of information in return. The detailed vehicle movement can be used to improve current traffic control algorithms (66; 98) such as speed harmonization (61; 65) and queue warning. An example of such improvement is the possibility to accurately detect shockwaves at the onset of formation, providing more flexibility in mitigating congestion.

Perturbations in traffic flow, a main factor for causing traffic congestion (5; 6), can be identified using the V2V/V2I technology. Unlike aggregate data sets generated by loop detectors, detailed vehicle trajectories can be used to measure disturbances caused by individual vehicles at
a much higher resolution. One way to estimate traffic disturbances is through estimating the speed variation among individual vehicles in a road segment (98). Previous studies suggested on the basis of both measurements and theory, using stochastic macroscopic models (99; 100) and microscopic models (101), that the widening of the speed distribution can serve as an early indicator of traffic breakdown. Kühne suggested using radar technology to measure the speed variation of vehicles over space as an early warning of traffic breakdown. The limitation in this method, however, is that it can only detect breakdown at the radar location and cannot track shockwave propagation across a segment.

Accordingly, this chapter presents a novel method to identify shockwave formation and track its propagation based on the speed distribution of individual vehicles available through connected vehicles technology. In addition, this chapter analyzes the impact of partial connectivity on shockwave identification and compares the accuracy of the proposed method to a wavelet transformation based method (97).

4.1 Data Description and Preprocessing

The Next Generation SIMulation (NGSIM) dataset was utilized to test the proposed method (102). The vehicle trajectories were collected on the southbound direction of US-101/Hollywood Freeway in Los Angeles, CA on June 15th, 2005, from 7:50AM to 8:35AM. The study segment contained five main lanes, one on-ramp, one off-ramp, and was approximately 2,100 feet long. The final dataset includes individual vehicles’ geographical coordinates, speeds, lane positions, and accelerations at every one-tenth of a second.
Montanino and Punzo’s (103) multistep corrections to NGSIM vehicle trajectories were applied as the first step for data preprocessing; however, no outliers were identified using the acceleration threshold suggested by the authors and hence interpolation was not necessary. The second step was to define spatial and temporal boundaries to remove errors on the aggregated estimates of traffic properties. The identified spatial boundary was the first 100-ft of the 2100-ft segment since mounted cameras might not have recorded all vehicles entering the segment.

4.2 Methodology

The 2000-ft segment (after identifying the spatial boundary), was divided into 200-ft sections to easily track traffic shockwaves (in this context, the propagation of speed drops over space and time). Different section lengths can be used for this analysis. While increasing the section length may reduce noise in aggregated traffic properties, it would reduce the accuracy of tracking shockwaves as the change in speed would be tracked over longer sections. Lu et al. (104) suggested using 500-ft (150-m) sections for aggregating NGSIM trajectories to estimate fundamental diagrams; however, that would be too long for the purpose of tracking shockwaves as the whole segment length is only 2000-ft. To easily track shockwaves over time, traffic properties were aggregated over 10-second time steps, an aggregation duration that was suggested in the Lu et al. (104) study as well.

Flow, mean speed, and density for each section at each time step were calculated using Edie’s generalized definitions for individual facilities (105; 106):

\[ q(A) = \frac{d(A)}{|A|} \]  

(4.1)
where \( q(A) \), \( k(A) \), \( v(A) \) denote flow, density, and mean speed for observed vehicles in section A, respectively, \( d(A) \) and \( t(A) \) represent the total distance traveled and total time spent by all vehicles in section A, respectively. Finally, \( |A| \) is the area covered by section A (200-ft section multiplied by a 10-sec interval).

The speed standard deviation (SSD) of individual vehicles in the main five lanes was calculated for each section at each time step. The following two steps were followed for this analysis: (1) calculating the average speed of each vehicle over 10 second period in a specific section; and (2) calculating the standard deviation of the average speeds of all individual vehicles in that section. This definition estimates spatial deviation in speeds among the vehicles in the five main lanes. The definition is slightly different from previous studies \( (99, 100) \) which used radar to measure individual vehicle speeds crossing a specific point over time. After estimating the traffic properties, the first and last time steps (10 sec) for each section were discarded as those time steps did not have complete trajectory information for the full 10 sec duration and flow/density values were underestimated.

The above calculations were performed for both full and partial connectivity cases. For the first case, all vehicles were assumed to be able to transmit their trajectories (location, speed, acceleration, etc.) at all times. However, full connectivity at the early stage of technology
deployment is highly unlikely. Therefore, in the second case, the mean speed and the speed standard deviation were calculated assuming partial connectivity for different market penetrations. In the partial connectivity case, a percentage of vehicles were assumed to be connected and able to transmit their trajectories. Hence, the calculations were performed for a sample of randomly selected vehicles. For example, at 30% market penetration, 30% of vehicles in the data set were randomly selected using the vehicle ID and the mean speed/standard deviation were calculated for this sample only for all sections at all time steps.

Note that the mean speed and the speed standard deviation estimates calculated for a sample of vehicle trajectories (partial connectivity) using the above-mentioned method can be very close to the true estimate (full connectivity) as discussed in the following section. However, estimating density and flow for a partially connected stream will result in values that are less than the true estimates. The reason is that only connected vehicles will be counted in this case instead of counting the actual number of vehicles, which will result in a lower density/flow estimate (unless model-based imputation techniques are invoked). The focus of this analysis, however, was to find out whether the speed standard deviation calculated at partial connectivity can still be used to detect traffic shockwaves. This will be discussed further in the following section.

4.3 Results and Discussion

4.3.1 Speed Standard Deviation as an Indicator of Shockwave Formation

The first step in the analysis is to explore the existence of patterns between the speed standard deviation (SSD) and shockwave formation in the dataset. To visualize those patterns, a time-space diagram was created for the mean speed and the SSD, as shown in FIGURE 8. The
direction of travel is from section 1 to 10. The figure indicates that shockwave formation is associated with a jump in the SSD. FIGURE 8(a), for example, shows six backward-propagating shockwaves that start in section 6 in the period 7:50 - 8:05 and a seventh shockwave that starts outside the segment of study. FIGURE 8(b) shows that each shockwave that started in section 6 is associated with a jump in the SSD that is also backward propagating. The other periods show the same pattern even though it is much noisier than the first period as the sections become more congested.
Another way to look at the relationship between speed and SSD is through a time-series graph. FIGURE 9 shows the mean speed changes over time compared to SSD changes during the time periods (a) 7:50AM – 8:05AM, and (b) 8:05AM – 8:20AM. The blue line refers to the mean speed on the left y-axis, while the orange line refers to the SSD on right y-axis. This figure shows that speed drops are associated with the peaks in SSD values that usually occur before the speed drops start. This correlation is clearer in the first time period (FIGURE 9(a)) than in the second one (FIGURE 9(b)) where most shockwaves in the second period start outside the study segment.
as seen in FIGURE 8(c). Therefore, FIGURE 8 and FIGURE 9 show that peaks in the SSD are better indicators of shockwaves that are starting to form than those which have already formed and are propagating.

FIGURE 9 Time-series graph for mean speed and speed standard deviation in section 4 during (a) period 7:50 - 8:05 and (b) period 8:05AM – 8:20AM

4.3.2 Propagation of Speed Standard Deviation Waves vs. Propagation of Shockwaves

While previous figures showed a pattern between SSD and mean speed, the figures do not show in which section SSD waves start and how they propagate backwards compared to shockwaves. FIGURE 10 shows the SSD waves (orange line) and speed shockwaves (blue line) in
sections 3 through 7. In section 7 (where speed shockwaves have not started yet), the SSD waves are starting to form signaling a reduction in speed in upstream sections. In section 6, the SSD waves show a clear signal while speed shockwaves are starting to form. In sections 5 through 3, the speed shockwaves and the SSD waves propagate backwards simultaneously at the same speed. Two conclusions can be drawn from this figure: (1) SSD waves start to form before speed shockwaves signaling a reduction in speed in upstream sections, and (2) the SSD waves propagate in conjunction with speed shockwaves at almost same speed. The propagation speed of SSD waves was estimated to be 10.2 mph which is the slope of the lines passing through the wave peaks. This is close to the 11.4 mph shockwave propagation speed that was estimated by Lu and Skabardonis (107) in their analysis of NGSIM trajectories.
FIGURE 10 Propagation of SSD spikes compared to the propagation of shockwaves during the period 7:50AM – 8:05AM. Direction of travel is from 3 – 7.

4.3.3 Flow – Density – Speed Standard Deviation relationship

FIGURE 11[a] shows flow-density diagram represented by the blue dots in addition to the speed standard deviation plotted against density on the right y-axis and represented by the orange dots. Each dot represents a value calculated for a section and a time step. The flow-density diagram (left y-axis in FIGURE 11[a]) shows that the traffic is operating close to capacity before the breakdown, as expected in the morning peak hour. The SSD – Density diagram (right y-axis in FIGURE 11[a]) shows that SSD increases around 80 – 90 vpmpl density value which is a typical
range for critical density before traffic breakdown (104). The increase in SSD and the scatter in the diagram reflects the instability of traffic flow during the transitioning phase from the uncongested to the congested state (98), as evident from the oscillations in FIGURE 8(a). After breakdown, SSD values drop reflecting less variation in speed of vehicles moving in the congested traffic regime. Note that this graph confirms Kühne’s findings in his study on freeway speed distribution (100). FIGURE 11[b] shows the empirical relationship between traffic flow and SSD. The trend line in the graph suggests that flow generally decreases at higher SSD values where traffic is usually unstable. However; it is important to note that the graph is highly scattered and deriving a definitive relationship requires further research.

4.3.4 Speed – Density – Speed Standard Deviation relationship

FIGURE 11[c] shows the speed-density graph represented by red dots in addition to the SSD plotted against density on the right y-axis, represented by the orange dots. As in the case of the above flow – density diagram, each dot represents a value calculated for a section and a time step. The speed-density diagram shows typical traffic behavior where the mean speed drops as the traffic density increases. Adding SSD to the graph shows around which speed the SSD starts to increase. In this case, SSD increases sharply right before 15-20 mph speed and drops after breakdown at around 10 mph. As in the case of previous, the graph is highly scattered and the abovementioned values are mere approximations. FIGURE 11[d] shows the relationship between mean speed and SSD. Similar to the trend in FIGURE 11[b], the traffic mean-speed decreases at higher SSD values. Both of the aforementioned trends suggest that SSD can be used as an indicator
of traffic disturbances. As the graphs show high scatter, however, deriving representative mathematical relationships requires further analysis and more data.

![Graphs showing Flow vs. Density vs. SSD, Mean Speed vs. Density vs. SSD, Mean Speed vs. Speed Standard Deviation for all periods (7:50AM – 8:35AM)](image)


### 4.3.5 Speed Standard Deviation Waves with Partial Connectivity

The above-mentioned graphs and discussions assume full connectivity of vehicles, in which case all vehicles are able to transmit their trajectories (speed, location, lane, etc.). However,
it is unlikely that traffic streams would be fully connected, at least not in the early stage of technology deployment. Therefore, the remaining sections will examine the SSD waves calculated using a sample of the provided trajectories, simulating scenarios where only part of the vehicles in the traffic stream are connected and able to transmit their trajectories. For example, in the 30% connectivity scenario, the SSD will be calculated using trajectories of 30% of the vehicles in the data set. Those vehicles are selected randomly using the vehicle ID. More importantly, the analysis will show whether the SSD waves detected for partially connected traffic can still be an indicator of shockwave formation.

FIGURE 12 shows SSD waves (colored lines) calculated for 10%, 20%, 30%, 70%, and 100% market penetrations. The mean speed in the graph (blue lines) is calculated at 100% connectivity to compare SSD waves to the actual speed shockwaves. The figure shows that at low market penetrations, specifically 10% and 20%, SSD could not be estimated for some time steps because there was not any connected vehicle detected (see disconnected lines in FIGURE 12). For market penetrations that are larger than 30%, SSD could be estimated for all time steps. Furthermore, the SSD – speed shockwave patterns are clear in the figure for medium to high market penetrations, specifically above 30%. This suggests that SSD waves can still be an indicator of shockwave formation at partial market penetrations above 30%. Note that choosing different seeds to randomly select the sample of vehicles may produce slightly different results. Moreover, it can be observed from the graph that at some time steps at very low market penetrations, SSD is either overestimated or underestimated, which means that low market penetrations can affect the quality of the SSD estimate.
As an indicator for the quality of SSD estimates, FIGURE 13 shows the number of vehicles used to estimate SSD at different market penetrations. Generally, a higher number of observations (the number of vehicles in this case) used to estimate a statistic leads to a higher quality estimate. At 10% and 30% market penetrations, the figure shows that the number of vehicles used to estimate SSD is below 15 (red lines), which suggests a low-quality estimate. At 50% market penetration, the quality increases with 15-25 observations used to estimate SSD for most time steps. Above
70% market penetration, the quality of the estimate is high with more than 25 observations used to estimate SSD for most time steps. These results support the pattern found in FIGURE 12.

FIGURE 13 Number of vehicles used to estimate the speed standard deviation at different market penetrations for section 2 in period 7:50AM – 8:05AM
4.3.6 Propagation of Speed Standard Deviation Waves at Different Market penetrations

FIGURE 14 shows the SSD waves estimated at 30% (green lines) and 50% (purple lines) market penetrations for sections 2 through 5 during the period 7:50AM – 8:05AM. The figure confirms that even for partially connected traffic streams, SSD waves can still serve as an indicator of shockwave formation because (1) the propagation of waves can be tracked over space and time, as seen in this figure, and (2) the pattern where a significant reduction in speed is associated with a peak in SSD is still evident.
FIGURE 14 Propagation of SSD waves at partial market penetrations
4.3.7 Comparison between SSD Waves and Speed Wavelet Transformation for Detecting Shockwaves

In a study by Zheng et al. (97), the authors proposed a Wavelet Transformation (WT) technique to identify bottleneck locations and activation times on freeways. Wavelet transformation is a time-frequency decomposition tool that extracts useful information from non-stationary time-series (97). In this method, the location of shockwaves and their formation time was extracted from transformed speed time-series. This is done by capturing local changes over time, a speed drop for example, through moving the wavelet’s location while squeezing and dilating the wavelet’s window. In their study, the authors used wavelet transformation to detect shockwave (bottleneck) formation on US-101 by transforming speed time-series of ten loop detectors. They identified the starting point of the shockwave by tracking the wavelet energy peaks that result from significant reductions in speed.

FIGURE 15 compares speed wavelet transformation (green line) to the SSD wave (orange line) estimated for section 3 during the period 7:50AM – 8:05AM. To estimate the wavelet energy distribution, the Mexican hat wavelet was used to transform the speed time-series and the distribution was estimated by averaging the wavelet energy over scales 1 through 64, as suggested by Zheng et al (97) and Talebpour et al (3). As seen in the figure, the wavelet energy is sensitive to every significant change in speed whether it is a decrease or an increase. To identify shockwaves in this case, cycles of energy peaks needs to be tracked. For example, the first peak identifies a shockwave (reduction of speed) and the second peak identifies the increase in speed that follows. Then, a third peak identifies a new shockwave. In the case of SSD waves, however, the peaks are
only sensitive to reductions in speed. Therefore, identifying shockwaves can be done more efficiently by tracking SSD peaks only.

Another feature to compare between speed WT and SSD is the responsiveness to speed reductions. FIGURE 16 shows the propagation of wavelet energy and SSD estimated for sections 3 through 6 during the period 7:50AM – 8:05AM. In section 6, where the shockwaves start to
form, the wavelet energy (green line) does not show any sign of disruption (except for the seventh shockwave which started outside the segment as seen in FIGURE 8a). The SSD values, on the other hand, are increasing, which indicates a reduction in speed that is about to occur in the following upstream section. In section 5, the speed reductions have started and the WT energy is responding to them. However, the full energy peak cycles (refers to an energy peak responding to a speed reduction and another peak responding to the recovery of speed after a shockwave is resolved, see FIGURE 15) have not developed yet. As for SSD values in section 5, they indicate the occurrence of shockwaves clearly. In section 4, the reductions in speed are more significant. Both speed WT and SSD are reflecting that. Again, however, the full peak cycles responding to speed reduction/increase have still not developed in section 4. Finally, in section 6, speed WT and SSD waves reflect the formation of shockwaves clearly. This graph suggests that while both
techniques can be used to detect shockwaves, SSD waves are more responsive to the reductions in speed than speed WT which can result in a higher detection accuracy.

FIGURE 16 Propagation of Wavelet Energy and SSD over Sections 3 – 6 during the Period 7:50AM – 8:05AM

4.4 Chapter Summary

Shockwave detection at the onset of congestion formation can contribute to more effective active traffic management techniques. Shockwave detection is traditionally done by tracking changes in speed and density of road segments using loop detectors. The accuracy of this method
is relatively low as it would depend on the number of loop detectors installed in a road segment and the spacing between them. Speed wavelet transformation is another method to track shockwaves by decomposing speed time-series. The technique can be used on speeds estimated by loop detectors, in which case it suffers from the same drawbacks as the previous method, or it can be used on speeds of individual vehicles. In the latter case, however, a high percentage of vehicles need to be able to transmit its trajectories in order to track the start of a shockwave accurately. This may be difficult in practice as the traffic flow is unlikely to be highly connected, at least not in the short run.

To address the above-mentioned limitations, this chapter explored the effectiveness of a method to detect shockwaves using the speed distribution of individual vehicles in a connected environment. The method is based on previous theoretical and empirical findings connecting the speed distribution to the onset of flow breakdown, though not specifically related to shock wave formation (99; 100). The main advantages of this method are: (1) shockwaves can be tracked accurately over small sections of a freeway (e.g., 200-ft segments were used in this analysis), and (2) shockwaves can be detected clearly for partially connected traffic streams. To test this method, vehicle trajectories from the NGSIM program were analyzed using the US 101 data set for a 2100-ft study segment in Los Angeles, CA. Two main cases were evaluated in this study: full connectivity and partial connectivity. In the case of full connectivity, it was assumed that all vehicles are able to transmit their detailed trajectories. In the partial connectivity case, only a percentage of vehicles were able to transmit their detailed trajectories and shockwaves were tracked using those trajectories only.
The results showed a consistent pattern where shockwave formation (significant reduction in speed propagating over space and time) is associated with a sharp increase in the speed standard deviation value (SSD) that usually occurs before the start of the shockwave development. Results also showed that SSD waves propagate in conjunction with speed shockwaves at almost the same speed, confirming that SSD waves can serve as an indicator to shockwave formation and propagation. These patterns were evident in both the full and partial connectivity cases. In terms of quality of the SSD estimate, the analysis showed that the quality improves at higher market penetrations as the number of vehicles (observations) used to calculate the SSD is higher. Compared to the wavelet transformation method, the analysis showed that SSD waves are more responsive to the reductions in speed than speed wavelet transformations, which can result in a higher shockwave detection accuracy. The higher accuracy in detecting shockwaves in terms of time and space can help improve the effectiveness of active traffic management techniques. It is worth mentioning however that this method was tested for a non-equilibrium traffic state where traffic was transitioning from an uncongested to a congested state. In an equilibrium (steady-state) traffic, a shockwave can occur with zero variance in speed and therefore may not be detected using this method.
5. MACHINE LEARNING APPROACH TO SHORT-TERM TRAFFIC CONGESTION PREDICTION IN A CONNECTED ENVIRONMENT

Traffic congestion is a complex phenomenon triggered by a combination of multiple interacting factors. \(5; 6\). Studies of real-world traffic breakdowns have identified three main factors as the main causes of traffic breakdowns \(5; 88\):

1) *High traffic loads*, where the traffic demand exceeds the average sustainable throughput on a road section;

2) *Local capacity reductions*, so-called “bottlenecks”, which may be permanent, such as on-ramps and off-ramps, or temporary such as traffic accidents or slow-moving vehicles.

3) *Disturbances caused by individual drivers*, which refer to temporary perturbations in the traffic flow, such as those caused by lane-change maneuvers, abrupt braking, or long-lasting overtaking maneuvers of trucks.

The first two factors, acting alone or jointly, generally lead to high density levels, at which the traffic flow becomes unstable, and less able to absorb perturbations due to the third factor, resulting in greater likelihood of flow breakdown \(89; 90\). High traffic loads and bottlenecks can be readily measured and analyzed using aggregated traffic data from traditional sensors (e.g., loop detectors and radars), or directly from the infrastructure configuration (for the second factor) such as on-ramps and lane drops. Disturbances in traffic flow, however, are difficult to observe using aggregated data \(88\) as they occur at the level of individual vehicles, and therefore are more difficult to control using techniques that rely on traditional sensor measurements.
Advances in vehicle wireless communications present new opportunities to measure traffic perturbations at the individual vehicle level and to improve the identification and prediction of congestion patterns (2; 4; 108-111). This new technology enables vehicles and infrastructure to exchange vehicle trajectory information (e.g., location, speed, and acceleration) through Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communications. The detailed trajectories can then be used to track finer patterns in vehicle movements (including the perturbations in traffic flow) and understand traffic flow dynamics in order to improve traffic prediction and control.

The key question is whether it is possible to utilize this detailed information to improve the identification and prediction of congestion formation. Machine learning is one way to take advantage of a large amount of information that can be generated by connected vehicles. These techniques have been studied in the literature for a wide range of transport applications, such as traffic operations, planning, and safety (112). One of the potential applications of machine learning that has not been extensively explored in the literature is predicting traffic congestion using vehicle trajectories. Some studies investigated congestion prediction with machine learning using other emerging data sources such as GPS records from mobile devices. For example, Thianniwet et al. (80), used a decision tree algorithm to classify road congestion by analyzing the movement patterns of vehicles collected through phone GPS. In another study by Pattara-Atikom and Peachavanish (81), the authors used neural networks to estimate traffic congestion from cell dwell time (the duration for which a cell is registered to a base station before moving to another base). Other studies conducted a time-series analysis to forecast traffic flow using neural networks (82; 83), or hybrid of multiple techniques (82; 84-87). For instance, in a study by Vlahogianni et al. (87), traffic patterns were recognized using clustering and then flow was forecasted using neural networks.
The main contribution of this chapter is to explore the capability of machine learning algorithms to predict short-term traffic congestion using vehicle trajectories available through connected vehicles. This is motivated by the various potential applications of these techniques for developing predictive traffic control algorithms in a connected environment. Logistic regression, random forests, and neural networks are used to develop two types of predictive models: (1) offline models which are calibrated based on historical data and are updated (re-trained) whenever significant changes occur in the system, such as changes/updates to the infrastructure, and (2) online models which are calibrated using historical data and updated regularly using real-time information on prevailing traffic conditions obtained through V2V/V2I communications. The remaining of this chapter provides a description of the vehicle trajectory data set used in the experiments, the methodology behind building the predictive models, and a discussion of the models’ prediction accuracy and their applications.

5.1 Data

The vehicle trajectory dataset used to build and test the predictive models was obtained from the Next Generation SIMulation (NGSIM) data for the US-101 Freeway in Los Angeles, California (102). The study segment is approximately 2100-ft long which contains five main lanes, one on-ramp, and one off-ramp. The trajectory dataset includes information on the vehicles’ speed, acceleration, location, and headways at a 0.1-second resolution.
5.2 Methodology

5.2.1 Estimating Traffic Properties

Estimating traffic properties follow the same process introduced in section 4.2. This chapter assumes a connected environment where vehicles share their detailed trajectories (i.e., location, speed, acceleration) through Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communications. In order to estimate the changes in traffic properties along the study segment of US-101, the 2100-ft segment was divided into ten 200-ft sections. Traffic flow, density, and mean speed were estimated for each section over 10-second time steps using Edie’s generalized definitions (105; 113):

\[ q(A) = \frac{d(A)}{|A|} \]  \hspace{1cm} (5.1)

\[ k(A) = \frac{t(A)}{|A|} \]  \hspace{1cm} (5.2)

\[ v(A) = \frac{d(A)}{t(A)} \]  \hspace{1cm} (5.3)

where \( q(A) \), \( k(A) \), \( v(A) \) denote flow, density, and mean speed for observed vehicles in section A, respectively. \( d(A) \) and \( t(A) \) represent the total distance traveled and total time spent by all vehicles in section A, respectively. Finally, \( |A| \) is the area covered by section A (200-ft section multiplied by a 10-sec interval).

The Speed Standard Deviation (SSD) of individual vehicles is the core of the predictive models in this study and was selected as a measure of perturbations in traffic flow at the individual
vehicle level. SSD was estimated for each section and time step as follows: (1) the average speed over the 10-second time steps was calculated for each vehicle passing a specific section, and (2) the standard deviation was estimated for the average speeds of all individual vehicles in that section. Next section discusses how SSD can be used to predict congestion.

5.2.2 Speed Standard Deviation as an Indicator for Congestion

As discussed in Chapter 1, disturbances caused by individual drivers constitute one of the main factors that cause traffic breakdown. The SSD of individual vehicles is one way to measure traffic perturbation (98) using detailed vehicle trajectories and is one of the main explanatory variables of the predictive models built in this study. SSD was particularly selected in this study because: (1) it has been suggested in the literature using macroscopic (100; 114) and microscopic (101) traffic models that the increase in speed variation (SSD) among individual vehicles is a prelude to flow breakdown, and (2) SSD can be estimated at a fine temporal and spatial scales (e.g. every 200-ft at 10-second intervals) using vehicle trajectories of fully or partially connected traffic stream.

5.2.3 Identifying Traffic States using K-means Clustering of the Fundamental Diagram

Identifying the traffic state from vehicle trajectories is the first step for building the predictive models in this study. The traffic state is used later as the dependent variable, which will be classified in the models. For simplicity, two traffic states were identified: uncongested (blue dots) and congested (red dots) as shown in FIGURE 17. A K-means clustering algorithm was used to identify the two states by specifying two clusters. The K-means clustering algorithm segments
data points into clusters (groups) where the total distances calculated from each point to its respective cluster center is minimized. As density and flow values are significantly different, the values were scaled to be between 0 and 1 before running the algorithm. Then, the density values were multiplied by a factor of 5 (arbitrarily chosen) to give density values a higher weight and force the algorithm to split the clusters based on density and define the critical density at which traffic breaks down. The critical density estimated using the K-means algorithm is around 80 vpmpl, which was also shown by a previous study using the NGSIM US-101 dataset (104).

FIGURE 17 Traffic states for the study segment of US-101
5.2.4 Temporally and Spatially Lagged Variables to Build Predictive Models

Mean speed and the SSD are the independent variables (explanatory variables) used in the models to predict the traffic state because: (1) both can be calculated accurately for a sample of connected vehicles, and (2) congestion is associated with high traffic perturbations (high SSD, see FIGURE 10) and low mean speed. While flow and density are also good indicators of the traffic state, both properties cannot be estimated accurately from vehicle trajectories of partially connected traffic as only connected vehicles will be counted.

The models developed in this study are temporally lagged models, which are a type of time-series models calibrated/trained to predict current values of the dependent variable, the traffic state, in this case, using lagged (past-values) of the explanatory variables. Thus, the models predict future values of the traffic state, our goal, when current values of the mean speed and SSD (the explanatory variables) are observed. The values of mean speed and SSD were lagged 10 and 20 seconds to calibrate models with 10-second and 20-second second prediction horizons which are intended for short-term prediction applications such as warning drivers of upcoming traffic slowdown, ramp metering, or speed harmonization.

In addition to temporally lagging SSD, the variable was spatially lagged 1 section downstream as the disruption in SSD waves starts downstream of the section where the shockwave (reduction in speed) begins to form. This is illustrated in FIGURE 10 where SSD waves in section 7 are signaling a shockwave formation that starts in section 6 and propagates upstream to sections 5 through 3. TABLE 3 provides a summary of the variables used in this study, their description, the market penetrations at which they were estimated, and the type of models they were used in.
Market penetration refers to the percentage of connected vehicles that are able to transmit their trajectories.
TABLE 3 Summary of the variables used in the predictive models

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Market Penetration Rates</th>
<th>Models</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent Variable</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Traffic State</td>
<td>Binary: the state of traffic whether congested or uncongested as identified using K-means clustering</td>
<td>100%</td>
<td>All</td>
</tr>
<tr>
<td><strong>Explanatory Variables</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lagged Mean Speed in Current Section (10 s)</td>
<td>Continuous: the average speed of individual vehicles in current section, lagged 10 seconds</td>
<td>30%, 50%, 100%</td>
<td>10 seconds predictive models</td>
</tr>
<tr>
<td>Lagged Mean Speed in Downstream Section (10)</td>
<td>Continuous: the average speed of individual vehicles in the next downstream section, lagged 10 seconds</td>
<td>30%, 50%, 100%</td>
<td>10 seconds predictive models</td>
</tr>
<tr>
<td>Lagged Speed Standard Deviation in Downstream Section (10)</td>
<td>Continuous: the speed standard deviation of individual vehicles in the next downstream section, lagged 10 seconds</td>
<td>30%, 50%, 100%</td>
<td>10 seconds predictive models</td>
</tr>
<tr>
<td>Lagged Mean Speed in Current Section (20 s)</td>
<td>Continuous: the average speed of individual vehicles in current section, lagged 20 seconds</td>
<td>30%, 50%, 100%</td>
<td>20 seconds predictive models</td>
</tr>
<tr>
<td>Lagged Mean Speed in Downstream Section (20 s)</td>
<td>Continuous: the average speed of individual vehicles in the next downstream section, lagged 20 seconds</td>
<td>30%, 50%, 100%</td>
<td>20 seconds predictive models</td>
</tr>
<tr>
<td>Lagged Speed Standard Deviation in Downstream Section (20)</td>
<td>Continuous: the speed standard deviation of individual vehicles in the next downstream section, lagged 20 seconds</td>
<td>30%, 50%, 100%</td>
<td>10 seconds predictive models</td>
</tr>
</tbody>
</table>
5.2.5 Offline Predictive Models

Offline predictive models are the first type of models calibrated and tested in this study. These models are built using historical data and updated whenever new data is available or as needed. The models are built using a training dataset that contains observed traffic states and a number of explanatory variables, such as mean speed and speed standard deviation in this case. Then, the models are used to predict the traffic state with the explanatory variables only.

The relevancy of data is key for the accuracy of offline models. For example, in order to predict the traffic state for a freeway on a summer day, the data used to build this model should cover a sample of summer days for the same freeway or a very similar one. If the data set used to build the model contains winter dates only, the accuracy will significantly drop as traffic dynamics are highly impacted by weather conditions. Ideally, in practice, historical data should cover a wide range of traffic states over different days of the week (weekdays, weekends) at different times (peak and off-peak hours) and weather conditions.

As the NGSIM US-101 data set is only available for a short time period (7:50AM – 8:35 AM) and for a single day, the whole data set was used to train the offline models and the accuracy was tested using K-fold cross-validation, which is suitable for small data sets when a separate test set is not available (115). The method will be discussed further in the section.

Three types of machine learning techniques were tested in this paper: (1) binary logistic regression, (2) random forests, and (3) artificial neural networks. The binary logistic regression model is a simple model that can help explain the relationships among the variables. Random
forests and artificial neural networks, on the other hand, are among the best machine learning classifiers (116) and are widely used for different transportation-related applications (112).

5.2.6 Online (Real-time) Predictive Models

Online models are the second type of models tested in this study. These models are built using historical data and updated (re-trained) regularly using real-time information on prevailing traffic conditions. The rate at which models are updated depends on how frequently traffic conditions are changing. During rush-hours for example, the models can be updated at a higher rate than during off-peak hours.

In this paper, the NGSIM dataset was used to simulate online models that are updated every five minutes. A high re-training rate was selected because the available NGSIM data covers a very short time period during the morning rush hour (45 minutes, 7:50AM – 8:35AM). We assumed that the historical data was similar to the first 10-min period (7:50AM – 8:00AM) of the total time period (7:50AM – 8:35AM). The model was initially built using data for the first 10-min period to predict the following 5-min interval (8:00AM – 8:05AM). Then, assuming that 5 minutes had elapsed, the models were retrained using the historical data (first 10 minutes) in addition to the new information received for the past 5 minutes. The model updating process was repeated until the whole time period was predicted (excluding the first 10 minutes). The updating process is illustrated in FIGURE 18. As in the case of offline models, three types of machine learning techniques were used to build online models: (1) binary logistic regression, (2) random forests, and (3) artificial neural networks.
FIGURE 18 Model updating process for online predictive models using NGSIM data

Note that training the model was assumed to be instantaneous. In practice, however, training can take a few seconds to few minutes depending on the size of the dataset, the machine learning technique used, and available computing resources. This needs to be taken into consideration when the updating process is designed. Another point to note is that the cumulative dataset up until the update point was used to retrain the models, which was not an issue since the available set is relatively small. A more efficient approach would be critical in practice when implementing these models for large datasets. One solution is to train them using a representative sample of the data.

5.2.7 Model Accuracy and K-fold Cross-Validation

The accuracy of the models tested in this paper is evaluated by three metrics:

a. Overall accuracy: the percentage of traffic states correctly predicted;

b. Congested state prediction accuracy: the percentage of the congested states correctly predicted; and
c. Uncongested state prediction accuracy: the percentage of the uncongested states correctly predicted.

In order to have a meaningful estimate of a predictive model’s accuracy, the model should be tested on a separate dataset (new observations) that was not used to train it (115). In the case of offline models, however, a separate data set was not available for testing. Therefore, a K-fold Cross-Validation technique was used to estimate the accuracy of the models.

K-fold cross-validation is a technique that is used to estimate the prediction accuracy (test error term) of a statistical learning method, such as the machine learning models used in this paper, by holding out a subset of the training data set from the fitting process, and then applying the models to predict those held out observations.

Using K-fold cross-validation to estimate the accuracy of the offline models, the data set was divided into 5 groups (folds) of equal size. The first group was treated as a validation group, and the remaining four groups were used to build the models. The models were then tested using observations of the first group, and the prediction accuracy (three metrics) was estimated. The process was repeated 5 times, a typical k-value in practice for this method (115), wherein each time a different group of observations was treated as a validation group. The k-fold Cross-Validation accuracy was then calculated by taking the average of the 5 different values estimated for each group. In the case of online models, the accuracy was estimated directly using the predicted observations since those models predict new observations that were not used to fit the models (see FIGURE 18).
5.2.8 Predictive Models with Partial Connectivity

In the case of partially connected traffic, mean speed and speed standard deviation (the explanatory variables) can be calculated using vehicle trajectories assuming that a minimum market penetration is required. This is illustrated in FIGURE 19, where the mean speed and standard deviation are estimated at different market penetration rates (i.e., 30%, 50%, and 100%). The graph shows that below 30% market penetration, mean speed and SSD could not be estimated for some time steps due to incomplete trajectories (there exist sections without any detected vehicles) or they are inaccurate. For higher market penetrations, mainly above 30%, the figure shows that mean speed and SSD estimates are very close to the true value (estimated at 100% connectivity) and therefore can be used to predict the traffic state.
FIGURE 19 Mean Speed and speed standard deviation estimated at different market penetrations for section 2 during the period 7:50AM – 8:05AM

Note that estimating flow and density directly for partially connected traffic will not be accurate, as only connected vehicles in the stream will be counted. Therefore, in order to use partial connectivity models, the trajectory data obtained from connected vehicles should be supported by extra information on density and flow to identify the observed traffic states (the dependent variable) and build the predictive models. This can be done through developing an inverse function to estimate density from mean speed, adjust the density/flow values directly estimated from trajectories, or collecting extra data on density using other methods such as cameras. This might
be easier to implement in the case of offline models as data collection is less frequent than online models.

5.2.9 Specifications of the Machine Learning Techniques

The three machine learning techniques were tested using the statistical software R\(^{(117)}\). The binary logistic regression was calibrated using the generalized linear model function in R. The cut-off probability beyond which the state was classified as congested was set to 50%. As for the random forest models, the randomForest package \(^{(118)}\) was used for calibration. Five hundred trees were selected to build the models as there had not been any improvement in the models’ accuracy when a higher number of trees were tested. Finally, the nnet package \(^{(119)}\) was used to calibrate neural network models using one hidden layer.

5.3 Results and Analysis

5.3.1 Offline Predictive Models with Full Connectivity

As discussed earlier, three types of machine learning techniques (binary logistic regression, random forest, and artificial neural networks) were used to build offline models. In this case, a fully connected environment was assumed in which all vehicles can exchange their trajectories using V2V/V2I communication technology. In addition, two prediction horizons were tested for these models: 10 and 20 seconds, which are intended for short-term prediction applications such as warning drivers of upcoming traffic slowdown, ramp metering, or speed harmonization.
TABLE 4 compares the different offline models built in this study with the assumption of full connectivity. The comparison is based on the three metrics defined in the Methodology section (overall accuracy, congested state prediction accuracy, and uncongested state prediction accuracy).

**TABLE 4 Comparison of Offline Models with Full Connectivity**

<table>
<thead>
<tr>
<th>Model</th>
<th>Overall Accuracy</th>
<th>Congested State Prediction Accuracy</th>
<th>Uncongested State Prediction Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic 10s</td>
<td>93%</td>
<td>96%</td>
<td>85%</td>
</tr>
<tr>
<td>Logistic 20s</td>
<td>91%</td>
<td>95%</td>
<td>79%</td>
</tr>
<tr>
<td>Random Forest 10s</td>
<td>93%</td>
<td>95%</td>
<td>85%</td>
</tr>
<tr>
<td>Random Forest 20s</td>
<td>90%</td>
<td>94%</td>
<td>77%</td>
</tr>
<tr>
<td>Neural Network 10s</td>
<td>89%</td>
<td>97%</td>
<td>68%</td>
</tr>
<tr>
<td>Neural Network 20s</td>
<td>90%</td>
<td>95%</td>
<td>78%</td>
</tr>
</tbody>
</table>

In terms of overall prediction accuracy, the table shows that the three techniques have high accuracy with values that range from 89% - 93% and that models with the shorter prediction horizon, 10 seconds, have higher accuracy than those with 20-second prediction horizon. Regarding congestion predicting, the three techniques performed well with accuracy that ranges from 94% - 97%. Additionally, the 10-second prediction horizon models performed marginally better (1% - 2% improvement) than the 20-second models. Finally, the prediction of the uncongested state was less accurate than the prediction of the congested state. This could be due to the noise in SSD waves which may imply congestion. The uncongested state prediction accuracy of the models was 68% - 85%. Out of the three machine learning techniques used, logistic regression and random forests performed slightly better than neural networks. In terms of training speed, logistic regression was the fastest. However, the differences were marginal as the dataset was small.
5.3.2 Offline Predictive Models with Partial Connectivity
TABLE 5 compares the three types of offline predictive models at low, medium, and full market penetrations (30%, 50%, and 100% respectively). The table shows that the predictive models perform well for partially connected traffic streams with an overall accuracy that ranges from 88% to 92% for both low and medium connectivity cases. This shows that full connectivity is not required to predict the traffic state using the above-mentioned techniques even though prediction accuracy improves when more vehicles in the traffic stream are connected. However, a minimum market penetration above 30% is required to estimate the values of mean speed and SSD as shown previously in FIGURE 19.
TABLE 5 Comparison of Offline Models with Partial Connectivity

<table>
<thead>
<tr>
<th>Model</th>
<th>Connectivity</th>
<th>Overall Accuracy</th>
<th>Congested State Prediction Accuracy</th>
<th>Uncongested State Prediction Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic 10s</td>
<td>30%</td>
<td>91%</td>
<td>96%</td>
<td>79%</td>
</tr>
<tr>
<td>Logistic 10s</td>
<td>50%</td>
<td>92%</td>
<td>95%</td>
<td>82%</td>
</tr>
<tr>
<td>Logistic 10s</td>
<td>100%</td>
<td>93%</td>
<td>96%</td>
<td>85%</td>
</tr>
<tr>
<td>Logistic 20s</td>
<td>30%</td>
<td>88%</td>
<td>94%</td>
<td>73%</td>
</tr>
<tr>
<td>Logistic 20s</td>
<td>50%</td>
<td>89%</td>
<td>94%</td>
<td>76%</td>
</tr>
<tr>
<td>Logistic 20s</td>
<td>100%</td>
<td>91%</td>
<td>95%</td>
<td>79%</td>
</tr>
<tr>
<td>Neural Network 10s</td>
<td>30%</td>
<td>91%</td>
<td>95%</td>
<td>80%</td>
</tr>
<tr>
<td>Neural Network 10s</td>
<td>50%</td>
<td>92%</td>
<td>95%</td>
<td>83%</td>
</tr>
<tr>
<td>Neural Network 10s</td>
<td>100%</td>
<td>89%</td>
<td>97%</td>
<td>68%</td>
</tr>
<tr>
<td>Neural Network 20s</td>
<td>30%</td>
<td>88%</td>
<td>94%</td>
<td>71%</td>
</tr>
<tr>
<td>Neural Network 20s</td>
<td>50%</td>
<td>89%</td>
<td>94%</td>
<td>75%</td>
</tr>
<tr>
<td>Neural Network 20s</td>
<td>100%</td>
<td>90%</td>
<td>95%</td>
<td>78%</td>
</tr>
<tr>
<td>Random Forest 10s</td>
<td>30%</td>
<td>91%</td>
<td>95%</td>
<td>80%</td>
</tr>
<tr>
<td>Random Forest 10s</td>
<td>50%</td>
<td>92%</td>
<td>95%</td>
<td>82%</td>
</tr>
<tr>
<td>Random Forest 10s</td>
<td>100%</td>
<td>93%</td>
<td>95%</td>
<td>85%</td>
</tr>
<tr>
<td>Random Forest 20s</td>
<td>30%</td>
<td>86%</td>
<td>92%</td>
<td>70%</td>
</tr>
<tr>
<td>Random Forest 20s</td>
<td>50%</td>
<td>88%</td>
<td>93%</td>
<td>73%</td>
</tr>
<tr>
<td>Random Forest 20s</td>
<td>100%</td>
<td>90%</td>
<td>94%</td>
<td>77%</td>
</tr>
</tbody>
</table>

As in the case of full connectivity, the models built assuming partial connectivity predicted the congested state better than the uncongested state due to the noise in SSD. The accuracy range for predicting the congested state was 92% - 96% while the accuracy range for predicting the uncongested state was 68% - 83%. In addition, 10-second prediction models, in general, performed better than 20-second prediction models, as in the case of full connectivity models.

5.3.3 Online Predictive Models

TABLE 6 compares the online predictive modes built using logistic regression, random forest, and neural networks techniques. The models used the first 10 minutes of the NGSIM data available as
a historical data and were updated every 5 minutes in real-time by adding the extra 5 minutes as new observations to the training dataset. The model updating process is illustrated in FIGURE 18.

In addition, the online models were built under the assumption of full connectivity. As mentioned in the methodology section, flow and density of traffic, which cannot be estimated at partial market penetrations, is required to identify the traffic states (the dependent variable) in new observations before adding them to the training dataset and update/retrain the models.

TABLE 6 Comparison of Online Predictive Model

<table>
<thead>
<tr>
<th>Model</th>
<th>Overall Accuracy</th>
<th>Congested State Prediction Accuracy</th>
<th>Uncongested State Prediction Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic 10s Online</td>
<td>91%</td>
<td>93%</td>
<td>87%</td>
</tr>
<tr>
<td>Logistic 20s Online</td>
<td>89%</td>
<td>94%</td>
<td>80%</td>
</tr>
<tr>
<td>Neural Network 10s Online</td>
<td>90%</td>
<td>93%</td>
<td>86%</td>
</tr>
<tr>
<td>Neural Network 20s Online</td>
<td>88%</td>
<td>94%</td>
<td>76%</td>
</tr>
<tr>
<td>Random Forest 10s Online</td>
<td>91%</td>
<td>94%</td>
<td>85%</td>
</tr>
<tr>
<td>Random Forest 20s Online</td>
<td>86%</td>
<td>93%</td>
<td>70%</td>
</tr>
</tbody>
</table>

The table shows that the overall accuracy of the predictive models is in the range of 86% - 91% which is slightly lower than the prediction accuracy of the offline models tested before. This also applies to the prediction accuracy of the congested state where the range is 93% to 94% compared to 94% - 97% in the case of offline models.

The results do not necessarily mean that online models perform worse than offline models. In fact, the results are influenced by the limited data set in this analysis. In the case of online models, the models were initially trained using only the first 10 minutes of the 45-minute period available in the data set and were used to predict the following 5 minutes. The models were then updated/retrained every 5 minutes while in each update new observations were added to the
training set. In other words, online models were built to improve over time to simulate a real-world application. Offline models, on the other hand, were built assuming that the whole 45 minutes of data was available as historical data and were all used to train the models once rather than build an initial model with a small dataset and improve it over time. Ideally, in practice, online models would use historical data as a basis for training and real-time information to improve the models.

5.4 Chapter Summary

Connected vehicles technology provides new opportunities to improve traffic congestion identification and prediction by generating detailed information on individual vehicle movements through V2V/V2I communications. Machine learning is one way to take advantage of the huge amount of information that can be generated by connected vehicles. Therefore, this chapter presented three machine learning techniques that can be used for short-term traffic congestion prediction using vehicle trajectories that would be available for connected vehicles. Two types of predictive models were developed: (1) offline models that are built using historical data only, and (2) online models that are updated in real-time.

The analysis shows that the prediction accuracy of the congested state by these models is 89% - 94% for the example data considered. Furthermore, the results show that congestion can be predicted accurately in the case of partially connected traffic streams, which is important in practice as traffic is unlikely to be fully connected, at least in the early deployment of this new technology.
The models have various safety and traffic performance applications. For instance, the models can be used to warn drivers of traffic slowdowns ahead to prevent potential accidents. In terms of traffic operations, the models can be integrated with traffic control algorithms to enhance their performance. For example, the models can be used to develop a predictive speed harmonization algorithm to predict the location of congestion and set the speed limit upstream of it (61), or to detect shock wave formation and provide queue warning information upstream (73). The prediction horizons of these models can be very short for some other applications, given that the prediction accuracy degrades for longer horizons using the current structure of the models. However, the analysis could be extended to predict longer horizons and multiple traffic states given vehicle trajectory data that covers a long duration and a wide range of states, though this matter would need to be empirically tested. Furthermore, the ability to predict changes and states that depart from what had been observed during the training period can degrade rapidly. Nonetheless, the high degree of predictability achieved for very short-term forecasts suggests that a productive approach would be to combine the advantages of machine learning techniques with those of approaches based on fundamental traffic concepts and theories.
6. CENTRALIZED CAV TRAFFIC MANAGEMENT APPLICATION – PREDICTIVE SPEED HARMONIZATION IN A CONNECTED ENVIRONMENT WITH CENTRALIZED SPEED CONTROL

Speed harmonization (SPDHRM - also known as variable speed limit in some applications) is an active traffic management strategy that has been widely used across the world to deal with congestion, incidents, or special events. This strategy changes the speed limit throughout the roadway segment of interest, based on prevailing traffic condition, roadway surface condition, and weather condition, to prevent or delay the onset of flow breakdown, mitigate congestion, and dampen shockwave formation. Current implementations of this strategy rely on fixed roadway sensors (e.g., loop detectors, radars, and/or video cameras) to assess the traffic condition in order to select the appropriate mitigation strategy (selection of speed limits throughout the target roadway segment). Moreover, changes in speed limit are displayed to drivers through variable message signs at fixed locations throughout the target segment. More sophisticated strategies implement lane-based advisory messages (e.g., Seattle, WA variable speed limit system).

Several speed harmonization models have been proposed and some have been implemented and tested in the field. An innovative example is SPECIALIST (120) that is focused on resolving moving shockwaves in the system. This algorithm utilizes the concepts from Kinematic Waves (96; 121) and tries to limit the inflow to the shockwave by reducing the speed limit. This algorithm was tested on a Dutch freeway and was able to increase the throughput by resolving the moving shockwaves (122; 123). In another study, Hegyi et al. (124) presented a predictive control model with a rolling-horizon approach to suppress shock waves by optimally coordinating the posted
speed limits in a speed harmonization system. Carlson et al. (125) introduced a cascaded control model to control the main traffic flow by adjusting the speed limits in a speed harmonization system. Chen and Ann (126) focused on fixed bottlenecks and developed a speed harmonization system based on shockwave theory. Their algorithm was able (in simulation studies) to increase throughput at fixed bottlenecks and reduce shockwave intensity (smooth speed transitions).

Even though current implementations of speed harmonization (i.e., utilizing fixed infrastructure sensors and communicating via variable roadway signs) may be able to delay flow breakdown and mitigate congestion, they face three key challenges. The first challenge is that fixed infrastructure sensors provide an incomplete picture of traffic flow dynamics throughout the segment of interest. For instance, shockwaves might not be captured until they pass a loop detector, which may be far from where they start. Such late detection can significantly affect the effectiveness of the speed harmonization strategy (127). The second key challenge is the limited set of scenarios in which speed control can be applied due to relying on fixed road sensors and signs, which significantly affects the strategy performance. Appropriate response to a detected shockwave is key to an effective and efficient speed harmonization system. However, communicating speed limit changes to the drivers at fixed locations via variable message signs can result in an ineffective response. For example, a traffic shockwave may pass a particular traffic sign location by the time a new speed limit is displayed. Furthermore, the updated speed limit, if not shown at the right time, may cause unnecessary slow-downs throughout the target segment. The third key challenge is the difficulty in predicting future traffic state utilizing data from fixed traffic sensors. As mentioned before, those sensors often collect aggregated data that are not
detailed enough to track finer traffic flow dynamics, which are essential for effective traffic control.

Recent advances in connectivity provide the opportunity to address the above challenges. Through V2I communications, vehicles broadcast information about their location, speed, acceleration, and direction to a central traffic management system, which in turn utilizes this information to capture finer traffic flow dynamics throughout the transportation system. Accordingly, shockwaves can be identified at the onset of formation \(^{(79, 127)}\), and more accurate prediction of future traffic states becomes possible \(^{(79)}\). Moreover, drivers can receive information from system operators about the transportation system status. Speed harmonization systems can significantly benefit from the introduction of connectivity, especially at high market penetration rates of connected vehicles. In such an environment, shockwaves can be detected more accurately and decisions about the speed limit can be directly transmitted to vehicles at any point along the segment of interest. Accordingly, many studies have investigated the development of speed harmonization systems in a connected driving environment. Grumert et al. \(^{(128)}\) proposed a cooperative speed harmonization system in a connected environment. They utilized V2I communications to provide individualized speed limit based on the distance from the incident location and showed that such a system can harmonize the traffic flow. In another study, Han et al. \(^{(129)}\) utilized connected vehicle technology to develop a speed harmonization system with the goal of improving throughput and reducing travel time at fixed bottlenecks. They tested three strategies: 1) one connected vehicle per lane without any variable message signs, 2) one connected vehicle per lane and utilizing variable message signs, and 3) multiple connected vehicles. They showed that speed harmonization could be effective even with a very limited number of connected
vehicles. Moreover, their findings suggest that any sudden change in speed limit could result in further shockwave formation and traffic flow instability.

The above and several other studies illustrated the effectiveness of utilizing connected vehicle technology for speed harmonization. However, achieving the full potential of speed harmonization system in a connected driving environment requires developing a reliable methodology for early shockwave detection, future traffic state prediction, and information dissemination strategies. For instance, Talebpour et al. (111) showed that signal interference and packet loss in a V2I communications system can significantly degrade the performance of the otherwise effective speed harmonization system.

Accordingly, this chapter puts forward a predictive speed harmonization system that utilizes the detailed vehicle trajectories broadcasted by CAVs through V2I communications and machine learning techniques to predict the location of traffic congestion. The system then determines the optimal speed limit to delay or prevent congestion and directly broadcasts those limits to CAVs upstream of the predicted congestion location. The main contribution of the system developed in this chapter is threefold: 1) it utilizes a novel shockwave detection method and machine learning techniques to predict traffic congestion anywhere on a freeway segment in real-time, 2) it utilizes V2I communications to collect detailed vehicle trajectories and to broadcast updated speed limits to connected vehicles without relying on fixed infrastructure sensors, and finally, 3) the system operates at low market penetrations of connected vehicles and in a mixed traffic environment with automated vehicles; both of these features are critical for practical implementation of CAV traffic control strategies.
6.1 Methodology

The predictive speed harmonization system proposed in this chapter relies on V2I communications and machine learning techniques to collect detailed trajectories from CAVs, estimate current traffic properties, predict future traffic state, and broadcast new speed limits to connected vehicles accordingly. It is different from traditional speed harmonization systems in three key areas:

1. It relies solely on connected vehicles to collect traffic information without any need to collect information from road sensors (e.g. loop detectors or radars), potentially reducing installation and maintenance cost of those sensors.

2. The system uses machine learning algorithms (data-driven) to predict traffic congestion and prevent/mitigate it. This data-driven approach improves the accuracy of prediction (up to 93% (79)) as it does not rely on strong assumptions about traffic behavior or road geometry.

3. The system accurately identifies the location of congestion anywhere on a freeway segment as it is not constrained by the location of infrastructure sensors.

The system consists of three main modules, illustrated in FIGURE 20: 1) Traffic Monitoring, 2) Congestion Prediction, and 3) Speed Control. The Traffic Monitoring module is responsible for collecting detailed traffic trajectories from connected and automated vehicles and utilizes an early shockwave detection method that tracks changes in the speed distribution of CAVs (79). The Congestion Prediction module identifies the location of the predicted traffic congestion anywhere on the targeted freeway segment (79). Finally, the Speed Control module determines
optimal speed limits (61; 130) based on prevailing traffic conditions and broadcasts those limits directly to CAVs upstream of congestion location. This system has three main design parameters: 1) prediction horizon, 2) broadcasting distance, 3) speed set. Prediction horizon is the duration (in seconds) before congestion is predicted to happen. Intuitively, prediction accuracy is lower at higher prediction horizons. Broadcasting distance is the distance between the predicted congestion location and the point at which CAVs receive updated speed limits before reaching congestion.
This module monitors the changes in traffic conditions and shockwave formations along the target freeway segment. As a central system, the module collects trajectories from all CAVs within a freeway segment of interest. It uses the early shockwave detection method (79) described in chapter 1 which tracks the speed distribution of individual CAVs over small road sections (e.g. 200 m) and time intervals (e.g. 10s). As discussed earlier in the chapter, the widening of the vehicles’ speed distribution measures traffic perturbations (98), a major contributor to traffic congestion (88), and is an early indicator of flow breakdown (79; 100; 114). The module monitors two main traffic properties: mean speed and the speed standard deviation (SSD). The mean speed is calculated using Edie’s generalized definition of speed (105; 106) according to the following formula:

\[ v(A) = \frac{d(A)}{t(A)} \quad (6.1) \]

where \( v(A) \) denotes the mean speed for observed vehicles in section A, \( d(A) \) is the total distance traveled by all vehicles in section A, and \( t(A) \) is the total time spent by all vehicles in section A. The SSD of individual vehicles is calculated in two steps: 1) the average speed of each vehicle is calculated over the monitoring time interval, 2) the SSD of the average speeds is calculated for each vehicle in the section. Both of the aforementioned properties are main inputs to the Congestion Prediction module.
6.1.2 Congestion Prediction

This module predicts the location of congestion formation within a short time horizon (10 – 30s) along the target freeway segment. The module utilizes the data-driven models discussed in chapter 5. It predicts traffic congestion at small sections (e.g. 200 m) of a freeway segment at small time steps (e.g. 10 seconds) (79). These models use the mean speed and the speed standard deviation (SSD) of individual vehicles estimated by the Traffic Monitoring module. The reason for choosing those properties only, without relying on traffic flow or density, is because mean speed and SSD can be estimated for a sample of connected vehicles (i.e. does not need traffic stream to fully connected). Therefore, the system can be used at the early deployment of CAV technology when traffic streams are partially connected.

The developed congestion models are temporally lagged models which are a type of time-series models trained to predict current and future values of the dependent variable using explanatory variable values observed in previous time steps. Therefore, by plugging in current values of independent variables, the model would predict future values of the dependent variable. In this case, the congestion prediction model was trained to predict the current traffic state (congested or not) based on mean speed and SSD values in previous time steps. In addition to temporally lagging SSD, the variable was spatially lagged one section downstream as widening of the speed distribution (increase in SSD) starts downstream of the section where congestion is starting to form. The general formulation of the lagged prediction models is as follows:

\[ y_{ts} = v_{(t-1)s} + v_{(t-1)(s+1)} + ssd_{(t-1)(s+1)} \]  
(6.2)
where $y_{ts}$ is the traffic state at time $t$ in section $s$, $v_{(t-1)s}$ and $v_{(t-1)(s+1)}$ are the lagged mean speeds in current and downstream sections respectively, $ssd_{(t-1)(s+1)}$ is lagged speed standard deviation in downstream section; $t$ denotes time step, $s$ denotes section number. A detailed description of the variables is provided in TABLE 7. Note that in this study, travel time index (TTI) was used to identify traffic congestion as suggested by Dong et al (89). Sensitivity analysis shows that this is a more robust approach to identify congestion than K-means clustering which was used in the original models by Elfar et al (79). The 2016 TTI index of Los Angeles, CA (1.7) was chosen to identify congestion in the following case studies. The LA index was chosen because the original models in chapter 5 were built using actual trajectory data collected in the same city. TTI for each freeway section is defined as follows:

$$TTI = \frac{Mean\ Speed^{-1}}{Free\ Flow\ Speed^{-1}}$$

(6.3)

**TABLE 7 summary of variables used in the predictive model**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Dependent Variable Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_{ts}$</td>
<td>Traffic State Binary: the state of traffic whether congested or uncongested as identified using the travel time index (TTI) (89) with a threshold above 2.</td>
</tr>
<tr>
<td>$v_{(t-1)s}$</td>
<td>Lagged Mean Speed in Current Section Continuous: the average speed of individual vehicles in the current section, lagged 10, 20, or 30 seconds</td>
</tr>
<tr>
<td>$v_{(t-1)(s+1)}$</td>
<td>Lagged Mean Speed in Downstream Section Continuous: the average speed of individual vehicles in the next downstream section, lagged 10, 20, or 30 seconds</td>
</tr>
<tr>
<td>$ssd_{(t-1)(s+1)}$</td>
<td>Lagged Speed Standard Deviation in Downstream Section Continuous: the speed standard deviation of individual vehicles in the next downstream section, lagged 10, 20, or 30 seconds</td>
</tr>
</tbody>
</table>

This study uses the random forests variant of the aforementioned models; the technique produces accurate traffic state predictions, needs relatively short training time, and doesn’t require data scaling or transformation (116). The random forest models previously tested by Elfar et al (79)
used actual vehicle trajectories collected through the Next Generation SIMulation (NGSIM) project in Los Angeles, CA (102). Due to the small size of the NGSIM data sample, however, the prediction model utilized in this study was calibrated using simulated vehicle trajectory data. Those vehicle trajectories were generated for 2-lane highway with one on-ramp at various demand levels (1000 - 2000 veh/mphl) using a simulation tool, discussed in section 6.2, that was developed for this study in the Python open-source programming language. The prediction accuracy of the aforementioned models is summarized in TABLE 8. Those results show that the prediction accuracy of the models is higher when a large data set is used to train the models (in the case of simulation).

**TABLE 8 Comparison of Traffic Congestion Prediction Models**

<table>
<thead>
<tr>
<th>Model</th>
<th>Prediction Horizon</th>
<th>Overall Accuracy</th>
<th>Congested State Prediction Accuracy</th>
<th>Un congested State Prediction Accuracy</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original Model - Chapter 5 (79)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logistic</td>
<td>10s</td>
<td>93%</td>
<td>96%</td>
<td>85%</td>
<td>NGSIM</td>
</tr>
<tr>
<td>Logistic</td>
<td>20s</td>
<td>91%</td>
<td>95%</td>
<td>79%</td>
<td>NGSIM</td>
</tr>
<tr>
<td>Random Forest</td>
<td>10s</td>
<td>93%</td>
<td>95%</td>
<td>85%</td>
<td>NGSIM</td>
</tr>
<tr>
<td>Random Forest</td>
<td>20s</td>
<td>90%</td>
<td>94%</td>
<td>77%</td>
<td>NGSIM</td>
</tr>
<tr>
<td>Neural Network</td>
<td>10s</td>
<td>89%</td>
<td>97%</td>
<td>68%</td>
<td>NGSIM</td>
</tr>
<tr>
<td>Neural Network</td>
<td>20s</td>
<td>90%</td>
<td>95%</td>
<td>78%</td>
<td>NGSIM</td>
</tr>
<tr>
<td>This study</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random Forest</td>
<td>10s</td>
<td>99%</td>
<td>95%</td>
<td>99%</td>
<td>Simulation</td>
</tr>
<tr>
<td>Random Forest</td>
<td>20s</td>
<td>98%</td>
<td>90%</td>
<td>99%</td>
<td>Simulation</td>
</tr>
<tr>
<td>Random Forest</td>
<td>30s</td>
<td>97%</td>
<td>87%</td>
<td>99%</td>
<td>Simulation</td>
</tr>
</tbody>
</table>

### 6.1.3 Speed Control

This module determines optimal speed limits to delay or prevent congestion once it is detected by the Congestion Prediction module and broadcasts the new limits to CAVs at a specific distance upstream of congestion location. For this application, an empirical speed decision tree introduced by Allaby et al (130) was adopted. The tree was tested in microscopic (61) and
mesoscopic (131) simulation applications in previous work. Through this decision tree, the speed is selected based on the mean speed of the freeway section at which congestion is predicted. The original Allaby tree was simplified, as illustrated in FIGURE 21, to include the congested case only as this when the speed control needs to be activated. Once a new speed limit is set, the module broadcasts it to CAVs at the specified broadcasting distance upstream of congestion location.

FIGURE 21 Speed decision tree for the SPDHRM system (61; 130; 131)

6.2 CAV Traffic Microsimulation Tool

To test the effectiveness of the SPDHRM system for multiple operational scenarios, the system’s logic was integrated into a CAV traffic microsimulation tool that was built using the open-programming language Python. The tool extends the CAV acceleration framework developed by Talebpour et al (73; 111; 132; 133) by introducing a CAV traffic monitoring component based on the methodology discussed in section 4.2 and a congestion prediction component based on the machine learning models developed in section 5.2. For the purpose of this study, two distinct driving behaviors were utilized in the CAV tool: connected and automated
driving behaviors. Below is a description of the acceleration models that captures those driving behaviors.

6.2.1 Modeling Connected Vehicles

Connectivity extends drivers’ perception of their surrounding environment beyond visual scanning of isolated drivers, leading to more responsive driving behavior (4). The information connected drivers receive can be related to the movement of vehicles ahead (e.g. speed and acceleration) which increases drivers’ awareness and perception. It can also be related to traffic conditions (e.g. congestion or weather conditions) that may impact driver’s strategic choices such as route choice or departure time. To capture the changes in driving behavior due to connectivity, the CAV simulation tool utilizes the Intelligent Driver Model (IDM) (6; 49).

IDM is a deterministic model that defines a vehicle’s acceleration as a continuous function of two parts: 1) the ratio of the vehicle’s current speed to its desired speed and 2) the ratio of the current spacing to the desired spacing. The first part controls the vehicle’s behavior in free flow traffic situations while the second part control its behavior in car-following mode. The model also considers perceptive parameters that can vary between drivers such as gap size and comfortable acceleration/deceleration. The IDM formulation is described below:

\[
a_{n, IDM}^{I}(s_n, v_n, \Delta v_n) = a_n \left[ 1 - \left( \frac{v_n}{v_0} \right)^{s_{n}} - \left( \frac{s^*(v_n, \Delta v_n)}{s_n} \right)^2 \right]
\]

(6.4)

\[
s^*(v_n, \Delta v_n) = s_0^n + T_n v_n + \frac{v_n \Delta v_n}{2 \sqrt{a_n b_n}}
\]

(6.5)
where $\delta_n$ is the free acceleration exponent, $T_n$ is the desired time gap, $a_n$ is the maximum acceleration, $b_n$ is the desired deceleration, $s_0^n$ is the jam distance, and $v_0^n$ is the desired speed. Those are parameters to be calibrated.

### 6.2.2 Modeling Automated Vehicles

Modeling automated vehicles considers two key features of the technology: 1) their ability to monitor other vehicles and entities in their surrounding using in-vehicle sensors, and 2) their ability to respond promptly to any changes in their environment. To capture those features, the CAV microsimulation tool employs a deterministic acceleration model developed by Talebpour and Mahmassani (4) that takes into account AV sensor range. The model formulations are presented below.

The maximum safe speed that is required for an automated vehicle to fully stop once its sensor detects an obstacle is calculated as follows:

\[
\Delta x_n = (x_{n-1} - x_n - l_{n-1}) + v_n \tau + \frac{v_{n-1}^2}{2a_{n-1}^{decc}} 
\]

(6.6)

\[
\Delta x = \min(\text{Sensor Detection Range}, \Delta x_n)
\]

(6.7)

\[
v_{max} = \sqrt{-2a_i^{decc}\Delta x}
\]

(6.8)

where $i$ and $i-1$ present the automated vehicle and its leader, respectively, $x_i$ is the location of vehicle $i$, $l_i$ is the length of vehicle $i$, $v_i$ is the speed of vehicle $i$, $\tau$ is the reaction time of vehicle $i$, and $a_i^{decc}$ is the maximum deceleration of vehicle $i$. FIGURE 22 visualizes the maximum safe speed threshold; any speed below the maximum safe speed curve is considered to be safe.
The automated vehicle movement is modeled as follows:

\[
a_i^d(t) = k_a a_{i-1}(t - \tau) + k_v(v_{i-1}(t - \tau) - v_i(t - \tau)) + k_d(s_i(t - \tau) - s_{ref})
\]  \hspace{1cm} (6.9)

where \(a_i^d\) is the acceleration of vehicle \(i\), \((k_a, k_v, k_d)\) are model parameters, \(s_i\) is the spacing and \(s_{ref}\) is the minimum between minimum distance \((s_{min})\), following distance based on the reaction time \((s_{system})\), and safe following distance \((s_{safe})\). In this study, minimum distance is set at 2.0 m and \(s_{system}\) and \(s_{safe}\) is calculated as follows:

\[
s_{safe} = \frac{v_{i-1}^2}{2} \left( \frac{1}{a_i^{decc}} - \frac{1}{a_{i-1}^{decc}} \right)
\]  \hspace{1cm} (6.10)

\[
s_{system} = v_i \tau
\]  \hspace{1cm} (6.11)

Finally, the acceleration of the automated vehicle is calculated using the following equation:

\[
a_i(t) = \min \left( a_i^d(t), k(v_{max} - v_i(t)) \right)
\]  \hspace{1cm} (6.12)

where \(k, k_a, k_v\) are model parameters to be calibrated. Based on the recommendations of Arem et al \((23)\), \(k = 1.0, k_a = 1.0, k_v = 0.58\), and \(k_d = 0.1\) were used in this study.
6.3 Results and Analysis

This section examines the impact of predictive speed harmonization on traffic flow performance and travel time for multiple operational scenarios of a two-lane highway segment (5 km) with an on-ramp merging traffic, illustrated in FIGURE 23. To do so, multiple performance measures were generated for the tested scenarios including fundamental (flow-density) diagrams, travel time distributions, speed distributions, and speed contours. Those graphs were generated for a freeway section 500 meters upstream of merging traffic to capture traffic congestion and stop-and-go behavior. The simulated scenarios below evaluate the effectiveness of the predictive speed harmonization system in fully and partially connected traffic streams. The scenarios also test the system’s performance and mixed traffic conditions (i.e. with automated vehicles). Note that in all of these scenarios, except for the partial connectivity one, all vehicles in the traffic stream are assumed to be connected.
FIGURE 23 Two-lane highway segment for testing multiple SPDHRM operational scenarios, main lanes volume 3000 veh/hr, ramp volume 500 veh/hr
6.3.1 Impact of the Predictive Speed Harmonization System on Traffic Performance

FIGURE 24 shows the impact of activating the speed harmonization system on traffic shockwave formation. The speed-contour graphs illustrate the temporal and spatial evolution of speed in the target segment. As discussed in the methodology, the segment in divided into small section (200 m) and speed in each section was measured at 10-second time steps. The dark blue lines illustrate the formation of traffic shockwaves, a significant drop in speed that propagates upstream over time and space. FIGURE 24(a) shows the base case scenario where SPDHRM is not activated. As seen in FIGURE 24, shockwaves start forming near the merging point in section 16 where traffic is disrupted due to the additional volume of traffic entering main lanes and the difference in speed of merging vehicles. Activating the system reduces the severity and length of traffic shockwaves, as seen in FIGURE 24(b), thereby improving the stability and safety of the traffic stream. This is because the SPDHRM system temporarily reduces traffic inflow upstream of congestion, thereby keeping flow below the road capacity and reducing the variation in vehicle speeds.
FIGURE 24 Impact of predictive SPDHRM on traffic shockwave formation (Centralized Control): (a) base case (b) active SPDHRM, broadcasting distance 1000m, prediction horizon 20s

FIGURE 25 depicts the impact of the SPDHRM system on traffic flow stability and breakdown. The flow-density diagram for the base case in FIGURE 25(a) shows that traffic breakdown, referring to the significant drop in flow, occurs around 80 veh/km density. This is caused by traffic perturbations generated at the merging point and volumes reaching segment capacity. Activating the SPDHRM system prevents traffic breakdown in this case as seen in FIGURE 25(b) where significant flow drops are eliminated. The system also improves the stability of traffic as indicated by less scatter in blue dots.
FIGURE 25 Impact of predictive SPDHRM on traffic flow stability and breakdown (Centralized Control): (a) base case (b) active SPDHRM, broadcasting distance 1000m, prediction horizon 20s

FIGURE 26 illustrates the impact of the SPDHRM system on the speed distribution of the target segment. The speed distribution of vehicles in FIGURE 26(b) shows that activating the system increases the overall average speed of the segment while also reducing the variation in speed compared to the base case in FIGURE 26(a). This indicates that the SPDHRM system improves speed homogeneity in the segment, the harmonization part of the system, as the continuous updating of speed limits helps vehicles drive at close speeds.
FIGURE 26 Impact of predictive SPDHRM on segment speed distribution (Centralized Control): (a) base case (b) active SPDHRM, broadcasting distance 1000m, prediction horizon 20s

FIGURE 27 shows the impact of the predictive SPDHRM system on vehicles travel time in the main lanes of the target segment. The travel time distribution in FIGURE 27(b) indicates that activating the system reduces the average travel time per vehicle compared to the base case in FIGURE 27(a). This is an expected reduction given that the overall speed of the segment is improved and the severity shockwaves are minimized.
FIGURE 27 Impact of predictive SPDHRM on vehicle travel time (Centralized Control): (a) base case (b) active SPDHRM, broadcasting distance 1000m, prediction horizon 20s
6.3.2 System Performance in Partial Connectivity Conditions

The previous simulations assume that all vehicles are able to broadcast their trajectories and receive advisory messages through the predictive speed harmonization system (full connectivity). In a real-world application, however, the traffic stream will be partially connected, at least at the early deployment of the technology. Therefore, this section tests the performance of the predictive control system for partially connected traffic streams to evaluate its applicability in the short term.

Connectivity affects the performance of the system in two major ways. Firstly, it impacts the amount of data received by the traffic management center (TMC) which affects the accuracy of congestion prediction. Secondly, connectivity impacts the percent of vehicles in the traffic stream that would be able to receive updated speed limits and therefore the overall traffic compliance with the new limit. As for prediction, results in section 5.3.2 show that the machine-learning-based predictive models require at least 30% connectivity to produce meaningful congestion predictions which generally have an accuracy above 90% (79). They also show that the marginal prediction improvement at high connectivity levels (market penetrations) is minimal. In this analysis, 30% connected traffic means that only 30 percent of vehicles, randomly selected, are able to send out their trajectories to the TMC and receive updated speed limits throughout the simulation.

FIGURE 28 shows the implications of partially connected traffic on the performance of the speed harmonization system for three cases: base (0% FIGURE 28(a, f)), low connectivity (40% FIGURE 28(b, e)), and high connectivity (80% FIGURE 28(c, f)). Results show the effectiveness of the speed harmonization system improves with higher connectivity as seen in less scatter of the
fundamental diagram and the shift towards lower values in the travel time distribution. The reason for this improvement is that more vehicles receive the updated speed limit which minimizes potential traffic disruption of partially slowing down upstream traffic. Those results confirm previous findings of the impact of connectivity on speed harmonization performance at both facility (61) and network level (131).
6.3.3 System Performance in Mixed Traffic Conditions

In a connected environment, traffic streams are likely going to be of mixed vehicle classes and driving behaviors especially with introduction of automated vehicles. The development of
automated vehicles is accelerating at an unprecedented rate; therefore, it is essential for any new traffic control strategy to account for this emerging driving behavior. To that end, this section tests the effectiveness of the predictive speed harmonization strategy at different market penetrations of automated vehicles: base (0% - no automation), low automation (30%), and high automation (70%).

Automated driving behavior is robotic in nature and therefore is inherently different from human driving behavior (2; 4; 108; 111; 134). It relies on in-vehicle sensors, high processing power, and programmed rules for environment perception rather than human judgement. This can make it safer and more stable than the human driving. The adopted acceleration framework in this study uses an automated vehicle (AV) car-following model introduced by Talebpour et al (135). This model is based on previous simulation studies by Van Arem et al (23) and Reece and Shafer (136). The AV model considers two main factors: (1) the ability of AVs to constantly monitor other vehicles in their vicinity, which can result in a deterministic behavior in dealing with other drivers’ behavior; and (2) their ability to react almost instantaneously to any changes in the driving environment.

FIGURE 29 shows the effectiveness of the proposed speed harmonization system in low automation traffic conditions (30% AV). To control for the impact of the more stable driving behavior of AVs, the figure plots the flow-density diagram and travel time distribution for low automation traffic case with and without activating the SPDHRM system. This would help show whether the improvement in traffic is due to the robotic driving behavior of AVs or the speed harmonization system. FIGURE 29(b) and FIGURE 29(e) show that AV driving behavior by itself
improves traffic performance by increasing traffic stability (less scatter in flow-density diagram) and lowering travel time compared to the base case in FIGURE 29(a, d). Activating the speed harmonization system, however, further improves traffic performance as seen in FIGURE 29(c, f). This improvement is due to controlling the speed of CAVs to improve congestion and the additional stability that AVs bring to traffic at low market penetrations.

FIGURE 30 shows the effectiveness of the predictive SPDHRM system in the high automation traffic case (70% AV). FIGURE 30(b, e) show that a high market penetration of AVs significantly improves traffic stability and reduce travel time without activating SPDHRM system. Moreover, FIGURE 30(c, f) show that the SPDHRM system was not activated at all as seen in the identical graphs to FIGURE 30(b, e). This is due to the dominant driving behavior of AVs, which substantially stabilize traffic and prevent congestion.
FIGURE 29 Effectiveness of the centralized SPDHRM system in mixed traffic environment – low automation: a, d) flow-density diagram and travel time distribution at 0% AV MPR with INACTIVE SPDHRM, b, e) flow-density diagram and travel time distribution at 30% AV MPR with INACTIVE SPDHRM, c, f) flow-density diagram and travel time distribution at 60% AV MPR with INACTIVE SPDHRM.
SPDHRM, c, f) flow-density diagram and travel time distribution at 30% AV MPR with ACTIVE SPDHRM, broadcasting distance 1000m, prediction horizon 20s

FIGURE 30 Effectiveness of the centralized SPDHRM system in mixed traffic environment – high automation: a, d) flow-density diagram and travel time distribution at 0% AV MPR with INACTIVE SPDHRM, b, e) flow-density diagram and travel time distribution at 70% AV MPR with INACTIVE SPDHRM, c, f) flow-density diagram and travel time distribution at 70% AV MPR with ACTIVE SPDHRM, broadcasting distance 1000m, prediction horizon 20s
6.4 Fine-tuning Design Parameters for Optimal Results

The centralized SPDHRM system has three main design parameters: 1) prediction horizon, 2) broadcasting distance, and 3) set of potential speed limits. Prediction horizon is defined as the time horizon (in seconds) over which traffic congestion is predicted to occur. For example, a 20 second prediction horizon means that the system will predict whether a specific roadway section will be congested in the next 20 seconds. Prediction horizon is an important parameter for two reasons: 1) it determines how early speed harmonization control is implemented and 2) it affects the accuracy of prediction. Activating speed harmonization too early can slow down traffic unnecessarily and therefore offset the expected benefits of the system. Inaccurate predictions, on the other hand, can activate speed harmonization falsely and therefore create new shockwaves.

To test the impact of prediction horizon on the system’s performance, three prediction horizon scenarios were tested as shown in TABLE 9: 10, 20, and 30 seconds. To control for the impact of broadcasting distance, it is set to 1000 m for all scenarios. The 20s prediction horizon case produces the best performance measures for the tested case study. This indicates that in the 10-second case, the control is applied too late after congestion was predicted (only 10 seconds) and therefore performed worse than the other two cases. The 30-second case improves the average travel time but does not improve speed homogeneity as much as the 20-second case. This could be due to the generation of other shockwaves caused by slowing down traffic earlier than necessary. Results for other broadcasting distances (500m, 1500m, and 2000m) follow a similar trend.
### TABLE 9 Impact of Prediction Horizon on SPDHRM System Performance

<table>
<thead>
<tr>
<th>Prediction Horizon (sec)</th>
<th>Average Travel Time (sec)</th>
<th>Average Speed (km/h)</th>
<th>StdDev Speed (km/h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>236</td>
<td>75</td>
<td>14</td>
</tr>
<tr>
<td>20</td>
<td>229</td>
<td>80</td>
<td>9</td>
</tr>
<tr>
<td>30</td>
<td>230</td>
<td>76</td>
<td>15</td>
</tr>
</tbody>
</table>

The second design parameter is the broadcasting distance which is defined as the distance between the predicted congestion location and a point upstream of that location at which CAVs receive updated speed limits (advisory messages) by the SPDHRM system. While the broadcasting distance itself is fixed, the location of the broadcasting area depends on the location of the predicted congestion which can occur anywhere on a freeway segment. This parameter is key to the predictive system's performance as it affects the distance over which incoming traffic is transitioning to a slower speed (lower flow) before reaching the congestion point and therefore the smoothness of transition. For example, slowing vehicles too far from the congestion location or too close may hinder the effectiveness of the strategy.

To test the impact of the broadcasting distance on the system performance, four broadcasting distance values were simulated as shown in TABLE 10: 500, 1000, 1500, and 2000 meters. The prediction horizon in those scenarios is set to 20 seconds to control for its impact. The 1000-m distance produced the best performance measures out of the four scenarios. The 500-m scenario indicates that the system does not offer enough distance for incoming traffic to transition smoothly to lower speeds generating compared to the 1000-m case. On the contrary, the 1500-m and 2000-m scenarios indicate that the system slows down traffic too far from congestion location and lead to less effective control. Simulations for other prediction horizons generated similar trend.
TABLE 10 Impact of Broadcasting Distance on SPDHRM System Performance

<table>
<thead>
<tr>
<th>Broadcasting Distance (m)</th>
<th>Average Travel Time (sec)</th>
<th>Average Speed (km/h)</th>
<th>StdDev Speed (km/h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>233</td>
<td>75</td>
<td>16</td>
</tr>
<tr>
<td>1000</td>
<td>229</td>
<td>80</td>
<td>9</td>
</tr>
<tr>
<td>1500</td>
<td>237</td>
<td>76</td>
<td>13</td>
</tr>
<tr>
<td>2000</td>
<td>235</td>
<td>77</td>
<td>13</td>
</tr>
</tbody>
</table>

The final design parameter is the set of potential speed limits that can be selected by the system to be broadcasted to CAVs once congestion is predicted. This parameter affects the rate at which incoming traffic is transitioning to a different speed and the overall speed of traffic. In previous simulations, three potential speeds were considered: 55, 75, and 90 km/h. Selecting the appropriate speed limit is based on a decision-tree that was developed through field experiments \(130\) and tested in mesoscopic \(131\) and microscopic \(61\) simulations. The decision-tree uses the average traffic speed of the congested section as a basis for the speed limit selection as seen in FIGURE 21.

In previous tests, the three parameters were assumed to be fixed throughout the duration of the strategy application. This setting, however, does not lead to optimal performance as the effectiveness of those parameters depends on the dynamic traffic conditions. For example, to mitigate severe traffic congestion, the SPDHRM system may require a longer distance to transition incoming flow (upstream of congestion) to lower speeds than light congestion. Therefore, the system needs a longer broadcasting distance to transmit updated speed limits. In addition, the decision-tree used for selecting the speed limit, while simple and practical, only offers three potential speeds to select from which may not be optimal for all traffic situations. To that end, the remaining of this section introduces an optimization-based formulation to select vehicle speed limits and broadcasting distance in order to maximize overall traffic speed, which in this context
minimizes overall travel time. The prediction horizon was set to 20 second since all previous tests showed that value to be optimal and to simplify the optimization problem.

6.4.1 Optimization Formulation for the Speed Control Module

The aim of this formulation is to maximize the overall traffic speed by varying the speed limits of vehicles upstream of traffic congestion once it is predicted to occur. The previous set-up of the Speed Control Module, discussed in section 6.1.3, selects the speed limit using an empirical decision tree that evaluates the speed of traffic in the congested section. This optimization formulation, however, evaluates a wider set of potential speed limits and selects the limit that maximizes traffic speed and by virtue mitigate congestion. In the most general formulation, a speed limit can be unique to each vehicle. This is only possible through CAV technology as speed limits can be broadcasted directly to vehicles. The mathematical formulation for speed limit selection is given below:

For each monitoring time-step (10 seconds), if congestion is predicted, update the speed limit of individual vehicles \( u_v^{m5} \) upstream of congestion location so that it:

Maximize:

\[
\max \sum_{t=t_0}^{t+t_{oh}} \sum_{v \in V} DT_{tv}(u_v^{m5})
\]

Subject to:

Speed limit boundaries:

\[ u_{min} \leq u_v^{m5} \leq u_{max}, \quad \forall v \in V^{us} \]

Multiple of 5 condition:
\[ u^{m5}_v = 5 \cdot u_v, \quad \forall v \in V^{us} \]  
(6.15)

Integer speed limit:

\[ u_v \text{ integer, } \quad \forall v \in V^{us} \]  
(6.16)

The objective function maximizes the travel distance \( DT_{tv}(u^{m5}_v) \) travelled by all vehicles \( V \) over an optimization horizon \( t_h \) by selecting an optimal speed limit (decision variable) \( u^{m5}_v \) for each vehicle. Note that maximizing distance traveled over a fixed time interval reflects maximizing traffic speed. The speed limit boundary condition ensures that the selected limits are within the minimum \( u_{min} \) and maximum \( u_{max} \) limits for the road segment of interest. The multiple of five condition limits the selected speed to multiples of 5 so that drivers can practically adhere to them. The integer speed condition ensures the speed is an integer for practicality as well. Note that the broadcasting distance in this formulation is implicitly optimized since all vehicles upstream of congested location receives an updated speed limit. Therefore, if a vehicle does not need to slow down, its optimal speed limit would be the same as the original speed limit of the road section (e.g. 100 km/h).

In order to solve the aforementioned formulation, the distance traveled function \( DT_{tv}(u^{m5}_v) \) needs to be estimated. However, the function cannot be estimated mathematically since it involves many intertwined vehicle interactions that can only be captured through traffic simulation such as car-following behavior, lane-changing behavior, multiple vehicle classes, and traffic control. Therefore, the distance traveled function in the objective function will be estimated using simulation, which will transform this problem into a simulation-based optimization problem.
The key limitation of simulation-based optimization problems is the computationally intensive and time consuming simulations that are associated with finding the optimal solution. In theory, solving this type of optimization problems requires simulating all possible solutions and choosing variables with best possible objective function value (137). This would be impractical for a real-time application since the above formulation requires simulating a finitely large number of potential solutions to reach an optimal solution. To put these numbers into perspective, assume that 100 vehicles are upstream of a congested location. If each vehicle can drive at 3 different speed limits, similar to the limit set in the currently used decision tree, solving the corresponding optimization problem would require simulating 300 potential combinations which would take days with moderate computing power making this approach unfit for real-time application.

To practically solve the optimization problem, the number of decision variables needs to be reduced significantly. Therefore, instead of solving for a unique speed limit per vehicle, the problem is reformulated to solve for one speed limit for all vehicles upstream of the predicted congestion location. Furthermore, the number of potential speed limits to be tested is finitely small given that potential limits need to be multiples of 5 and needs to be within the minimum and maximum speed limits of the road segment of interest. The modified formulation introduces broadcasting distance explicitly to find the best point upstream of congestion at which vehicles needs to receive the updated speed limit. The potential distances are also finite as seen in the formulation below.

Maximize:
\[
\max \sum_{t=t_0}^{t+t_{oh}} \sum_{v \in V} DT_{tv}(u, d)
\]  

(6.17)

Subject to:

Speed limit set:

\[u \in U = \{u_{\text{min}}, (u_{\text{min}} + 5), \ldots, u_{\text{max}}\}\]  

(6.18)

Broadcasting distance set:

\[d \in D = \{500, 1000, 1500\}\]  

(6.19)

Similar to the original formulation, the objective function maximizes the distance traveled \(DT_{tv}(u, d)\) by all vehicles \(V\) over an optimization horizon \(t_{oh}\) by selecting a speed limit \(u\) and broadcasting distance \(d\) (decision variables) for all vehicles upstream of traffic congestion. The speed limit set and the broadcasting distance set constraints limit the optimal solution to a predefined set of feasible speeds \(U\) and distances \(D\) to be able to solve this problem in real-time. The following section tests the effectiveness of the optimization-based SPDHRM system and compares it to the previous system using speed decision tree.

6.4.2 Performance Comparison: Optimization-based versus Decision-Tree Speed Control

This section compares the performance of the optimization-based control strategy with the performance of the speed decision-tree control strategy. To do so, a 2-lane highway with one on-ramp, illustrated in FIGURE 23 was simulated for both strategies. The pre-defined speed limits for the optimization-based control were \(U = \{55, 60, 65, \ldots, 100\} \text{ km/h}\). The set of pre-defined broadcasting distances were \(D = \{500, 1000, 1500\} \text{ m}\). As for the decision-tree control strategy, the tree introduced in section 6.1.3 was used. Note that the speed limit values in the decision tree,
FIGURE 21, are a subset of the speed limit set $U$. The optimization horizon was set to 30 seconds (3 times the monitoring time-step).

FIGURE 31 shows traffic shockwave formations when the SPDHRM system is activated for both control strategies. Using the optimization-based approach further reduces the severity and length of traffic shockwaves as seen in FIGURE 31(b) compared with the decision-tree approach in FIGURE 31(a). This improvement is mainly due to two reasons. First, the optimization-based strategy selects an optimal speed limit from a wider set of speeds which leads to a smoother transition of incoming (upstream) flow to mitigate congestion. Second, the optimization strategy selects an optimal broadcasting distance dynamically which ensures that vehicles have enough distance to transition to different speed.

FIGURE 31 Traffic shockwave formation patterns: a) speed decision-tree, b) optimization-based speed control

FIGURE 32 compares traffic stability for both control strategies. The flow-density diagram in FIGURE 32(b) show that the optimization-based control improves the stability of traffic as evident from the less scatter in the diagram compared to FIGURE 32(a). The improvement in
stability is due to the smoother transition of upstream traffic in the case optimization-based control. 

FIGURE 33 compares the overall traffic speed for both control cases. The speed distribution in FIGURE 33(b) shows that optimization-based approach increases the overall traffic speed and reduces its variation. This is expected given that the objective function in the optimization formulation maximizes traffic speed. Finally, FIGURE 34 shows the impact of both control strategies on vehicle travel time in main lanes. The travel time distribution in FIGURE 34(b) shows that the optimization-based approach leads to lower average travel time for vehicles. This is also expected given that maximizing speed leads to lower travel times.
FIGURE 32 Flow-density diagrams: a) speed decision-tree, b) optimization-based speed control

FIGURE 33 Speed distributions: a) speed decision-tree, b) optimization-based speed control

FIGURE 34 Travel time distributions: a) speed decision-tree, b) optimization-based speed control
6.4.3 Applying optimization-based control in a real-world scenario

There are two key challenges to consider when applying the aforementioned formulation in a real-world scenario. The first challenge pertains to adding an extra layer of prediction as result of using traffic microsimulation to estimate the distance traveled by vehicles (objective function) in the formulation. This would increase the overall likelihood of prediction errors in the system since in this case the error can be due to predicting a wrong traffic state or due to discrepancies between traffic simulation and actual traffic behavior. To overcome this challenge, the traffic simulation tool to be used for this application needs to capture the traffic dynamics of mixed traffic conditions (including CAVs). This can be done by using state-of-the-art car-following and lane changing models, such as the acceleration framework developed by Talebpour et al (111), and calibrating those models to the road segment of interest. Traffic simulation can also be supported by reinforced learning techniques to improve its performance (138).

The second key challenge of applying optimization-based control is the high computational effort required to solve the optimization formulation whenever traffic congestion is predicted. Estimating the total distance traveled by vehicles whenever a speed limit is updated can only be done through simulation; therefore, solving the optimization formulation requires simulating the entire potential set of solutions and selecting the one with the best outcome. For example, if the SPDHRM system has a set if 10 speed limits and 3 broadcasting distances to select from, it would simulate a total of 30 potential scenarios each time congestion is predicted to select the optimal limit and distance. This may take a considerable amount of time that may deem real-time application infeasible. This challenge can be overcome in multiple ways. One way would be to simulate potential scenarios in parallel, rather than in sequence, which can significantly save execution time since these scenarios are independent. Another way is to optimize the traffic simulation tool itself for speed so that it can evaluate the potential scenarios faster. Finally, the set of potential speeds and broadcasting distance can be reduced in size to reduce the number of scenarios to be evaluated.
TABLE 11 compares system performance and simulation time of multiple optimization-based control setups to the decision-tree control approach. As seen in the table, the best performer out of the 5 setups is case 5, which has the most extensive set of potential speed limits and broadcasting distances to evaluate. This setup, however, is also the most computationally intensive as seen from the simulation running time (9 h). Reducing the number of potential speed limits to evaluate, as in the case of setup 3, reduces the computational time significantly (1.5 h) while maintaining a performance that is close to the best case (225 vs 221 s travel time). This indicates that with careful selection of a small set of potential speed limits to test, the execution time can be reduced significantly without sacrificing performance considerably.
TABLE 11 Control strategy comparison: decision-tree vs optimization based

<table>
<thead>
<tr>
<th>No.</th>
<th>Control Strategy</th>
<th>Speed Range (km/h)</th>
<th>Broadcasting Distance (m)</th>
<th>Evaluated Scenarios (No.)</th>
<th>Average Travel Time (sec)</th>
<th>Average Speed (km/h)</th>
<th>Simulation-time (45 min run)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Base (INACTIVE)</td>
<td>NA</td>
<td>NA</td>
<td>0</td>
<td>241</td>
<td>73</td>
<td>~45 min</td>
</tr>
<tr>
<td>2</td>
<td>Decision-tree</td>
<td>100, 90, 75, 55</td>
<td>1000</td>
<td>0</td>
<td>229</td>
<td>80</td>
<td>~45 min</td>
</tr>
<tr>
<td>3</td>
<td>Optimization-based</td>
<td>100, 90, 80, 70, 60</td>
<td>500</td>
<td>5</td>
<td>225</td>
<td>82</td>
<td>~1.5 h</td>
</tr>
<tr>
<td>4</td>
<td>Optimization-based</td>
<td>100, 90, 80, 70, 60</td>
<td>1000</td>
<td>5</td>
<td>231</td>
<td>87</td>
<td>~1.5 h</td>
</tr>
<tr>
<td>5</td>
<td>Optimization-based</td>
<td>100, 95, 90, 55, 55, 1000, 1500</td>
<td>500, 1000, 1500</td>
<td>30</td>
<td>221</td>
<td>85</td>
<td>~9 h</td>
</tr>
</tbody>
</table>

6.5 Chapter Summary

This chapter presented a predictive speed harmonization system that relies on a traffic management center to collect information from CAVs within a road segment of interest, predict traffic congestion location, and broadcast updated speed limits to CAVs upstream of that location to mitigate congestion. The system consists of three main modules: 1) Traffic Monitoring, 2) Congestion Prediction, and 3) Speed Control. The Traffic Monitoring module collects detailed traffic trajectories from CAVs and utilizes the early shockwave detection method that tracks changes in speed distribution of CAVs. The Congestion Prediction module identifies the location of the predicted traffic congestion anywhere on the targeted freeway segment. Finally, the Speed Control module determines optimal speed limits to mitigate congestion based on prevailing traffic conditions and broadcasts those limits directly to CAVs upstream of congestion location.
Analysis of multiple operational scenarios showed that activating the predictive speed harmonization system reduces the severity and length of traffic shockwaves and improves the overall stability of traffic. It also improves the overall traffic speed and reduces the average travel time per vehicle. The system performance analysis of partially connected traffic stream indicates that the system’s performance improves at higher connectivity levels. This improvement is due to the higher number of CAV trajectories collected by the system and the more accurate congestion prediction at higher connectivity levels. The system effectiveness analysis in mixed traffic conditions showed that the speed harmonization system improves traffic performance at low market penetration of AVs. At medium to high AV market penetrations, however, the robotic driving behavior of AVs dominates the traffic stream; therefore, it significantly stabilizes traffic and entirely prevents breakdown.
7. DECENTRALIZED CAV TRAFFIC MANAGEMENT APPLICATION

– PREDICTIVE SPEED HARMONIZATION IN A CONNECTED ENVIRONMENT WITH DECENTRALIZED SPEED CONTROL

A decentralized speed harmonization system relies on an embedded control logic within CAVs to collect detailed vehicle trajectories shared by other CAVs, predict traffic congestion ahead using vehicle-specific models, and adjust their speed to mitigate congestion. The core of this strategy is similar to the centralized SPDHRM system discussed in chapter 6; vehicles approaching a congested section (bottleneck) reduce their speeds temporarily to avoid getting stuck in congestion for extended periods. Unlike a centralized system where the control logic is executed by a traffic management center, however, the control logic is executed in a decentralized context where CAVs lower their speeds based on the information they receive from other vehicles ahead and the specific control logic programmed in it. The prediction model and the speed selection logic can be unique for each vehicle or a group of vehicles. Historically, such system would not be possible due to limitations in computing power and communications, however, advancement in both areas allow CAVs to share information with each other in real-time through V2V communications and decide upon an optimal speed that would minimize congestion without the need to any V2I communication devices. This can potentially save costs related to deploying V2I communications and traffic management centers.
7.1 Methodology

The decentralized version of the predictive speed harmonization system shares the main features that differentiate the centralized system from traditional set-ups, but are applied in a decentralized context as follows:

1) The decentralized system relies on CAVs only to collect traffic information through V2V communications without any need to collect information from road sensors (e.g. loop detectors or radars), potentially reducing installation and maintenance cost of those sensors.

2) The system uses machine learning algorithms embedded in CAVs to predict traffic congestion and prevent/mitigate it. While those models do not rely on strong assumptions about traffic behavior, their accuracy depends on the number of CAVs that are able to communicate with each other.

3) The system accurately identifies the location of congestion ahead of a CAV anywhere in a freeway segment as it is not constrained by the location of infrastructure sensors.

The decentralized system, illustrated in FIGURE 35, consists of three main modules integrated into individual CAVs: 1) Traffic Monitoring, 2) Congestion Prediction, and 3) Speed Control. The Traffic Monitoring module collects detailed traffic trajectories shared by other CAVs within the vehicle’s communication range and utilizes the early shockwave detection method, discussed in chapter 1, to track changes in speed distribution of downstream CAVs (79). The Congestion Prediction module identifies the location of the predicted traffic congestion.
downstream of the target vehicle anywhere on the targeted freeway segment (79). Finally, the Speed Control module determines optimal vehicle speed (61; 130) based on downstream traffic conditions. This system has two main design parameters: prediction horizon and broadcasting distance. The former is the duration (in seconds) before congestion is predicted to happen. Intuitively, prediction accuracy is lower at higher prediction horizons. Broadcasting distance is the distance between the predicted congestion location and the point at which a vehicle starts to slow down.
7.1.1 Traffic Monitoring

This module monitors the changes in traffic conditions and shockwave formations downstream of a target vehicle. As a decentralized system, the module only collects trajectories from CAVs within the target vehicle’s communication range. This is more significant in the case of dedicated short-range communications (DSRC) where range is very limited (e.g. 100 – 1000m). For other technologies such as 4G or 5G networks, the range can be virtually unlimited depending on network coverage. In such cases, other factors such as communication delay and road segment size would be more important. The number of trajectories collected also depend on the number of
CAVs within the communication range. The module uses the early shockwave detection method (79) described in chapter 1 to track the speed distribution of individual CAVs over small road sections (e.g. 200-ft) and time intervals (e.g. 10s), ahead of the target vehicle. As discussed earlier, tracking speed distribution is an effective way to predict congestion since the widening of vehicles’ speed distribution serves as an early indicator to traffic breakdown (79; 100; 114).

Since the evaluation distance and location differ for each vehicle, the road can either be divided into smaller sections relative to each vehicle’s location, or it can be pre-programmed as small sections into vehicles in the form of a map. The former method would require extra computational time as road discretization would need to be applied frequently due to the continuous movement of the vehicle. The latter case, which is adopted in this chapter’s simulations, assumes that the road configuration is programmed into the vehicle itself and is divided into small sections (links). This is similar to the operation of current CAVs under development that use mapping information in addition to sensors for environment perception (e.g. Google’s self-driving car).

The module monitors two main traffic properties: mean speed and the speed standard deviation (SSD), both of which are inputs to the Congestion Prediction module. Those are calculated using the same formulas discussed in section 6.1.1.

7.1.2 Congestion Prediction

This module predicts the location of congestion formation within a short time horizon (10 – 30s) downstream of a target vehicle. The module utilizes the data-driven models discussed in chapter
5. Unlike the centralized case, however, those models can differ from vehicle to vehicle based on the logic programmed into them by manufacturers. Vehicles can have one standard model that they all share or multiple models (or fleets) that are unique to their manufacturers. In all cases, the models predict traffic congestion at small road sections (e.g. 200m) and at small time steps (e.g. 10 seconds) (79) downstream of a target vehicle. These models use the mean speed and the speed standard deviation (SSD) of individual vehicles estimated by the Traffic Monitoring module. The reason for choosing those properties only, without relying on traffic flow or density, is because mean speed and SSD can be estimated for a sample of connected vehicles (i.e. does not need traffic stream to fully connected). Therefore, the system can be used at the early deployment of CAV technology when traffic streams are partially connected. Since this is a decentralized control case with potentially multiple prediction models, the prediction models’ accuracy is significantly affected by the number of CAVs using the same model. For example, a fleet of OEM A vehicles would only be able to connect to other OEM A vehicles and therefore the prediction accuracy of their models would depend on the number of OEM A vehicles on the road. This accuracy would improve though if OEM A, in this case, allows cross-communication with other OEM vehicles to obtain more trajectory information. The formulation of the prediction models is similar to the centralized version discussed in detail in section 6.1.2.

7.1.3 Speed Control

This module determines the optimal speed limit to delay or prevent the onset of flow breakdown once it is detected by the Congestion Prediction module, and applies that speed if the vehicle is within a specific distance from congestion. This module uses the same empirical speed decision
tree \([61; 130; 131]\) that was used in the centralized control version although it is applied individually by vehicles. The tree is discussed in detail in section 6.1.3.

### 7.2 Results and Analysis

This section explores the impact of the decentralized SPDHRM on traffic flow performance and travel time for multiple operational scenarios of a two-lane highway segment (5 km) with an on-ramp merging traffic, illustrated in FIGURE 23. To do so, multiple performance measures were generated for the tested scenarios including fundamental (flow-density) diagrams, travel time distributions, speed distributions, and speed contours. Those graphs were generated for a freeway section 500 meters upstream of merging traffic to capture traffic congestion and stop-and-go behavior. The simulated scenarios below evaluate the effectiveness of the SPDHRM system in fully and partially connected traffic streams.
7.2.1 Impact of the Decentralized Predictive Speed Harmonization System on Traffic Performance

FIGURE 36 shows the impact of activating the decentralized SPDHRM system on traffic shockwave formation. The speed-contour graphs illustrate the temporal and spatial evolution of speed in the target segment. As discussed in the methodology, the 5 km freeway segment is divided into small section (200 m) and the speed in each section was measured at 10-second time steps. The small sections are assumed to be preprogrammed in vehicles so that they would not have to discretize the road in real-time. The dark blue lines illustrate the formation of traffic shockwaves, a significant drop in speed that propagates upstream over time and space. FIGURE 36(a) shows that the base case scenario where SPDHRM is not activated. As seen in the figure, shockwaves start forming near the merging point in section 16. By activating the SPDHRM system, the severity and length of shockwaves are reduced as shown in FIGURE 36(b), improving the stability and safety of the traffic stream. Unlike the centralized case shown FIGURE 24(b), decentralized speed control generates light shockwaves near the section where vehicles start to slow down. This is due to the variation in location at which vehicles start to slow down since each vehicle detects congestion at a slightly different location depending on its communication range and the number of CAVs downstream. However, the overall improvement in performance is significant as seen in the speed and travel time distributions below.
FIGURE 36 Impact of predictive SPDHRM on traffic shockwave formation (Decentralized Control): (a) base case (b) active SPDHRM, broadcasting distance 1000m, prediction horizon 20s, Single Prediction Model

FIGURE 37 shows the impact of the decentralized SPDHRM system on traffic flow stability and breakdown. The flow-density diagram for the base case in FIGURE 37(a) shows that traffic breakdown, refers to the significant drop in flow, occurs around 80 veh/km density. This is caused by traffic perturbations generated at the merging point and volumes reaching segment capacity. Similar to the centralized system, activating the decentralized SPDHRM system prevents traffic breakdown and improves traffic stability. This is seen in FIGURE 37(b) where significant drops in traffic are eliminated and the values are less scattered.
FIGURE 37 Impact of predictive SPDHRM on traffic flow stability and breakdown (Decentralized Control): (a) base case (b) active SPDHRM, broadcasting distance 1000m, prediction horizon 20s, Single Prediction Model

FIGURE 38 shows the impact of the SPDHRM system on the speed distribution of the target segment. The speed distribution of vehicles in FIGURE 38(b) shows that activating the system increases the overall average speed of the segment while also reducing the variation in speed compared to the base case in FIGURE 38(a). As expected, the SPDHRM system improves speed homogeneity in the segment as the continuous updating of speed limits helps vehicles drive at close speeds.
FIGURE 38 Impact of predictive SPDHRM on segment speed distribution (Decentralized Control): (a) base case (b) active SPDHRM, broadcasting distance 1000m, prediction horizon 20s, Single Prediction Model

FIGURE 39 shows the impact of the predictive SPDHRM system on the vehicles’ travel time in the main lanes of the target segment. The travel time distribution in FIGURE 39(b) indicates that activating the system reduces the average travel time per vehicle compared to the base case in FIGURE 39(a). This is an expected reduction given that the overall speed of the segment is improved, and the severity of shockwaves is minimized.
FIGURE 39 Impact of predictive SPDHRM on vehicle travel time (Decentralized Control): (a) base case (b) active SPDHRM, broadcasting distance 1000m, prediction horizon 20s, Single Prediction Model
7.2.2 System Performance in Partial Connectivity Conditions

Connectivity affects the performance of the decentralized SPDHRM system in three major ways: 1) it impacts the number of vehicles that are communicating and therefore the range at which CAVs can detect congestion, 2) it impacts the congestion prediction accuracy given that only CAVs are able to share their trajectories with nearby CAVs, and 3) it impacts the percent of vehicles in a traffic stream that would slow down to mitigate congestion. The machine learning models are capable of accurately predicting congestion at low connectivity levels as seen from the results in section 5.3.2. The detection range, however, is more sensitive to connectivity levels as it affects the range at which vehicles can detect congestion. This is more apparent in the case of multiple fleets where CAVs only communicate with other CAVs in the same fleet. In this case, the detection range would be significantly lower which would impact the effectiveness of the system.

FIGURE 40 shows the implications of partially connected traffic on the performance of the speed harmonization system of a single CAV fleet where a single prediction model is used by all vehicles. The figure shows the fundamental diagram and travel time distribution for three scenarios: base (0% FIGURE 40 (a, f)), low connectivity (40% FIGURE 40 (b, e)), and high connectivity (80% FIGURE 40 (c, f)). Results show that the effectiveness of the speed harmonization system improves with higher connectivity. This can be seen by the higher stability of traffic (less scatter in the flow-density diagrams) and the shift towards lower values in the travel time distribution. The reason for this improvement is that more vehicles are able to share their information and apply the control logic. This also confirms some of the previous findings of the
impact of connectivity on speed harmonization performance at both facility (61) and network level (131).

FIGURE 41 shows the implications of partially connected traffic on the performance of the speed harmonization system of multiple CAV fleets where each fleet has its own prediction model and communicates only to vehicles within that fleet. The figure compares the travel time distribution of vehicles in the case of three CAV fleets to the case of a single CAV fleet. For both low connectivity and high connectivity cases, the decentralized system with multiple CAV fleets performed worse than a system with a single fleet. This can be seen in the shift towards lower values in the travel distributions in the case of multiple fleets FIGURE 41(a, c) compared to those of a single fleet FIGURE 41(b, d). This lower performance is mainly due to two reasons: 1) the lower range at which CAVs in multiple fleets can detect congestion given that vehicle within a fleet can only communicate with other vehicles in the same fleet, and 2) the lower congestion prediction accuracy of CAVs in multiple fleet as those only receive information from vehicles in the same fleet and therefore lower the quality of estimated traffic properties. From a policy perspective, these results indicate that a successful implementation of a decentralized control system requires open communication between different vehicle classes (e.g. OEMs or fleets) in order to collect more vehicle trajectories and improve congestion prediction accuracy. It would also require standardizing prediction model development by OEMs to reduce the variety in congestion predictions and improve control consistency. This can be done by, for example, specifying the properties of the machine learning model to be used by vehicles and the specific variables to include.
FIGURE 40 Impact of partial connectivity on the SPDHRM’s performance (Decentralized Control): a – c) flow density diagrams for 0%, 40%, and 80% CV market penetration respectively, d-f) travel time distributions for 0%, 40%, and 80% CV market penetration respectively, broadcasting distance 1000m, prediction horizon 20s, single prediction model
Highly automated vehicle systems are likely to be part of the traffic mix in the near future. They exhibit fundamentally different driving behavior and therefore they need to be captured while testing CAV traffic control strategies such. To that end, this section examines the performance of the decentralized speed harmonization system in partially automated traffic streams at low (30\%) and high (70\%) market penetrations. As mentioned before, the automated vehicle models used in the following simulations are adopted from Talebpour et al.’s acceleration framework (\textsuperscript{4}).
FIGURE 42 shows the effectiveness of the decentralized speed harmonization system in low automation traffic conditions (30% AV). The figure controls for the effect of AV driving behavior by plotting the flow-density diagram and time distribution with and without activating the decentralized SPDHRM system. Therefore, the plots show whether the performance improvement, if any, is caused by the AV driving behavior or activating the predictive control strategy itself. FIGURE 42(b) and FIGURE 42(e) show that AV driving behavior by itself improves traffic performance by increasing traffic stability (less scatter in flow-density diagram) and lowering travel time compared to the base case in FIGURE 42(a, d). Activating the decentralized SPDHRM system does not produce significant changes in traffic stability beyond the improvement incurred by automated vehicles as seen in FIGURE 42(c). The system, however, reduces the average travel time as shown in FIGURE 42(f). This improvement is due to controlling the speeds of CAVs in addition to extra stability that AVs brings to traffic.

FIGURE 43 shows the effectiveness of the decentralized SPDHRM system in the high automation traffic case (70% AV). FIGURE 43(b, e) show that a high market penetration of AVs significantly improves traffic stability and reduce travel time without activating SPDHRM system. Moreover, FIGURE 43(c, f) show that the decentralized SPDHRM system was not activated at all as seen in the identical graphs to FIGURE 43(b, e). This is due to the dominating driving behavior of AVs which totally stabilize traffic and prevent congestion.
FIGURE 42 Effectiveness of the centralized SPDHRM system in mixed traffic environment – low automation: a, d) flow-density diagram and travel time distribution at 0% AV MPR with INACTIVE SPDHRM, b, e) flow-density diagram and travel time distribution at 30% AV MPR with INACTIVE SPDHRM, c, f) flow-density diagram and travel time distribution at 30% AV MPR with ACTIVE SPDHRM.
SPDHRM, c, f) flow-density diagram and travel time distribution at 30% AV MPR with ACTIVE SPDHRM, broadcasting distance 1000m, prediction horizon 20s

FIGURE 43 Effectiveness of the decentralized SPDHRM system in mixed traffic environment – high automation: a, d) flow-density diagram and travel time distribution at 0% AV MPR with INACTIVE SPDHRM, b, e) flow-density diagram and travel time distribution at 70% AV MPR with INACTIVE SPDHRM, c, f) flow-density diagram and travel time distribution at 70% AV MPR with ACTIVE SPDHRM, broadcasting distance 1000m, prediction horizon 20s
7.3 Centralized vs. Decentralized Speed Harmonization Systems

The two types of the predictive speed harmonization system introduced in this dissertation – centralized and decentralized – share the same main components: traffic monitoring, congestion prediction, and speed control. They differ however in the way each of those components operates and the wireless telecommunication technology that they predominately rely on. For both systems, the traffic monitoring component is responsible for collecting detailed vehicle trajectories broadcasted by CAV vehicles in real time. The centralized SPDHRM system relies on a traffic management center and V2I communications to collect and process CAV data in order to estimate fundamental traffic properties such as the traffic mean speed and the speed standard deviation. The decentralized SPDHRM system, on the other, operates at the individual vehicle level where the vehicles themselves share their information between each other through V2V technology and estimate traffic conditions downstream of their location in real time. In this case, vehicles would have a smaller point of view of the traffic ahead compared to central management center as the amount of data collected by individual vehicles is generally smaller than a traffic management center.

The congestion prediction component of both systems is responsible for predicting traffic congestion within a small time horizon (10-30s) using machine learning tools. The centralized version of the system predicts congestion anywhere on a freeway segment using the data collected from CAV vehicles. In this case, the system would use a unified prediction model. The decentralized version does the prediction at individual vehicle level; each CAV uses the information shared by surrounding vehicles and an embedded model to predict traffic congestion
downstream of its location within a short time horizon. In this case, the prediction model itself can differ from vehicle to vehicle depending on the manufacturer (or fleet).

The speed control component of both strategies determines the optimal speed at which vehicles should drive in order to mitigate congestion once predicted. The centralized system determines the optimal speed based on the information received from all CAVs within a freeway segment of interest. The decentralized system, however, allows individual vehicles to determine their optimal speeds based on the information they receive from surrounding CAVs, usually downstream vehicles. Therefore, a centralized system would have a more complete view of prevailing traffic conditions and could apply control more uniformly.

In terms of performance, activating either the centralized or decentralized version of the SPDHRM system results in reducing severity of traffic shockwaves, improving the stability of traffic, increasing average traffic speed, and reducing travel time. Those improvements are generally more significant at higher market penetrations of CAVs. The centralized version outperforms the decentralized one in two main cases: 1) when the centralized system implements optimization-based speed control or 2) when the decentralized system operates with multiple CAV fleets that do not cross-communicate. In the first case, a simulation-based optimization problem is solved in real-time to choose the optimal speed limit to minimize congestion. This approach assures that CAVs are slowing down at an optimal rate and only when it is necessary to slowdown. In some cases, the optimal action, even when congestion is detected, is to actually maintain the original speed limit. Solving the optimization problem is only applicable in the centralized version because it requires a full view of the whole freeway segment of interest, typically achieved by a
central traffic management center. In the second case where multiple CAV fleets do not share information between each other (e.g. fleets of different OEMs), the amount of information collected by CAVs is small and therefore their congestion prediction is less effective. A decentralized approach can also lead to small variation in activating the speed reduction which can create minor disruptions in traffic as seen in FIGURE 36(b).

When choosing which version is more practical to implement, four key factors can be considered: 1) the system’s implementation requirements, 2) associated costs, 3) integration with other active strategies, and 4) impact of traffic automation on performance. While both systems share the same core components, their essential implementation requirements differ substantially. A centralized system relies on V2I wireless telecommunications, which requires the installation of infrastructure point-of-contact devices to be able to receive information from CAVs - particularly in the case of DSRC technology. The centralized system also requires a traffic management center (or multiple ones depending on network size) to collect and analyze the information broadcasted by CAVs. Maintaining both the infrastructure point-of-contact devices and the traffic management center would usually fall under the authority of a public agency such as the state or city department of transportation. A decentralized system, on the other hand, relies on V2V wireless telecommunications to share and collect information from surrounding CAVs. Therefore, it requires integrating the system’s logic into individual vehicles without the need for external devices or traffic management centers. For a successful implementation of the system, as seen from the results in section 7.2.2, the data collection process and the congestion prediction model need to be standardized across individual vehicles. Otherwise, activating the system can vary
considerably from one vehicle to another which reduces the effectiveness of the system, more substantially at low CAV market penetrations.

In terms of the costs associated with each system, the decentralized version may have a lower overall cost. A decentralized SPDHRM system does not require installing or maintaining a traffic management center or any infrastructure devices that may be required for V2I telecommunications. The costs of the chips required implement the system can be spread out over individual CAV vehicles which will have the computational power required to collect and process data shared by other vehicles. Moreover, current smartphones, which have a processing power on par with personal computers, can be connected to the car to perform most of the data analysis, potentially lowering implementation costs further.

Integrating speed harmonization with other active traffic management or safety strategies can be easier in the case of a centralized system. Most active traffic strategies require a traffic management center that collects data from road sensors and therefore the same center can be used for the predictive speed harmonization strategy. For example, a single TMC can be used to implement a network-level route optimization and speed harmonization to minimize overall congestion. SPDHRM can also be integrated with a queue warning system or a collision warning system to improve overall safety. This integration of multiple strategies can lower the overall costs associated with implementing SPDHRM as it would share the same resources with other strategies. A decentralized version can also be coupled with other strategies, such as queue warning, depending on how much processing an individual vehicle can handle.
The last main factor to consider when choosing whether to implement either of the two versions is the impact of traffic automation on traffic performance. As discussed in the literature review and some of the simulation experiments above, automated vehicles are expected to have a more stable driving behavior that would significantly improve traffic performance and safety. At low market penetrations of AVs, activating SPDHRM further improves the stability of traffic and travel time beyond the improvements caused by the automated driving behavior itself, as discussed in section 6.3.3 and 7.2.3. However, the SPDHRM system has little to no impact at high market penetration of AVs as the highly automated driving totally stabilizes traffic and prevents congestion. In other words, the benefits of SPDHRM system can be significant only in the short-term when the market penetration AVs are low. Therefore, the short-term benefits of the predictive SPDHRM system, assuming that AVs will achieve mass market penetration in the long-term, may not justify the high investment in a centralized SPDHRM system by itself. It may be cost-efficient if the system is integrated with other active management or safety strategies such as routing or collision warning. Implementing a decentralized system, on the other hand, can be worth it given that it has less costly requirements and that car manufacturers are heavily invested in AV technology which basically has all the capabilities required to implement a decentralized SPDHRM system (sensors, machine learning, wireless communications, etc.). TABLE 12 summarizes the comparison between the two variation of the speed harmonization system.
### TABLE 12 Centralized vs. Decentralized Speed Harmonization System

<table>
<thead>
<tr>
<th>Performance Factors</th>
<th>Centralized SPDHRM</th>
<th>Decentralized SPDHRM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic Shockwaves</td>
<td>Reduces intensity and number of shockwaves</td>
<td>Reduces intensity and number of shockwaves but can create minor disruptions</td>
</tr>
<tr>
<td>Traffic Stability</td>
<td>Improves stability – less scatter in fundamental diagram</td>
<td>Improves stability – less scatter in fundamental diagram</td>
</tr>
<tr>
<td>Traffic Speed</td>
<td>Improves mean speed (up to 85 km/h) and reduces variation (up to 7 km/h)</td>
<td>Improves mean speed (up to 79 km/h) and reduces variation (up to 6 km/h)</td>
</tr>
<tr>
<td>Travel Time</td>
<td>Reduces average travel time (up to 221 sec)</td>
<td>Reduces average travel time (up to 220 sec)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Implementation Factors</th>
<th>Centralized SPDHRM</th>
<th>Decentralized SPDHRM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implementation Requirements</td>
<td>Requires additional infrastructure development Higher</td>
<td>Relies on in-vehicle technology only Lower</td>
</tr>
<tr>
<td>Integration with Other Strategies</td>
<td>Easy integration Improves performance in low traffic automation conditions – no impact in high traffic automation conditions</td>
<td>Requires further modification Improves performance in low traffic automation conditions – no impact in high traffic automation conditions</td>
</tr>
</tbody>
</table>

- The performance measures shown in this table are for a 2-lane highway (5km). Please refer to the results sections in previous chapters for detailed results

### 7.4 Chapter Summary

This chapter presented a decentralized speed harmonization system that relies on individual CAVs to collect data through communicating with each other, predict traffic congestion using vehicle-specific models, and adjust their speeds in order to mitigate congestion. The decentralized system consists of three main modules embedded in individual CAVs: 1) Traffic Monitoring, 2) Congestion Prediction, and 3) Speed Control. The Traffic Monitoring module collects detailed traffic trajectories shared by other CAVs within the vehicle’s communication range and utilizes the early shockwave detection method, discussed in chapter 1, to track changes in speed distribution of downstream CAVs. The Congestion Prediction module identifies the location of the predicted traffic congestion downstream of the target vehicle anywhere on the targeted freeway
segment. The Speed Control module determines optimal vehicle speed based on downstream traffic conditions.

Results showed that activating the decentralized system reduces the severity of traffic shockwaves, improves stability of traffic, increases overall traffic speed, and reduces travel time. Unlike the centralized system, the decentralized version generated minor shockwaves due to the non-uniformity in predicting traffic congestion and applying control among individual vehicles. However, the overall traffic performance improved for both control strategies. The analysis also showed that having multiple prediction models (fleet-based models) reduces the effectiveness of the strategy. Therefore, a successful application of the decentralized system requires standardization of data collection among vehicles and the ability to communicate with vehicles from other fleets to improve prediction range and accuracy. Compared to the decentralized version discussed in the previous chapter, implementing a decentralized strategy can be more cost effective and practical. The strategy does not require any more capabilities beyond what CAV would generally have and many vehicle OEMs are already invested in the technology.
8. DECENTRALIZED CAV TRAFFIC MANAGEMENT APPLICATION

– TRUCK PLATOONING IN MIXED TRAFFIC ENVIRONMENT

Truck platooning is a special application of the decentralized CAV traffic control strategies that can potentially improve traffic performance, in addition to reducing freight costs. It links multiple trucks in a convoy using wireless telecommunications and automated control systems. The algorithm assigns one truck to be designated as the leader and others in the convoy adjust their speeds to that of the leading vehicle, following each other at short distances. Truck platooning can rely on CACC technology to control longitudinal movement by adjusting the speeds of following trucks or, as envisioned for future systems, use fully automated driving functions (longitudinal and lateral) that rely on wireless telecommunications and vehicle sensors.

Truck platooning has potential safety, mobility, and sustainability benefits. In terms of safety, platooning can improve the reaction of connected trucks over individual trucks as the platooned vehicles can adjust their speeds to that of the leading vehicle, minimizing the likelihood of an accident due to a slower reaction. This can improve further if the leading truck is automated and uses multiple sensors to detect traffic around it. As for mobility benefits, truck platooning can improve the efficiency of operating trucks on the road, which can improve the traffic state. In addition, improved speeds translate in reduced travel times for goods in transit, thereby lowering logistics costs for freight shippers. Finally, moving at short distances reduces air drag between trucks significantly and therefore improves fuel consumption, lowers emissions, and reduces overall costs.
As the percentage of trucks on some interstate highways can be significant, trucks can be a significant factor in the performance of mixed traffic as a whole. This chapter focuses on modeling truck platooning in a mixed connected environment to evaluate the overall operational performance. To do this, a modeling framework of automated truck platooning developed by PATH at the University of California, Berkeley (139) was adopted and integrated into the CAV traffic microsimulation tool introduced in section 6.2. The truck platooning modeling framework is discussed in the methodology section below.

8.1 Methodology

This chapter adopts a truck platooning modeling framework developed by PATH at the University of California, Berkeley (139; 140). This framework has three distinct driving behaviors for automated trucks: 1) cruise control, 2) adaptive cruise control, and 3) cooperative adaptive cruise control. Below is a description of each of those behaviors. The model parameters were calibrated using experimental data collected by researchers at PATH (139; 140).

8.1.1 Cruise Control (CC)

The cruise control car-following behavior captures automated truck driving in free flow conditions (no leading vehicle) or when a time gap to a leading vehicle is above a certain threshold (2.5 seconds). Through this control logic, an automated truck maintains a desired speed—in this case, the freeway speed limit. The car following formula is as follows:

\[ a_{Aut}(t) = k_p \left[ v_{ref}(t - 1) - v(t - 1) \right] \]  

(8.1)

where \( k_p \) is a model parameter (0.3907)
8.1.2 Adaptive Cruise Control (ACC)

The adaptive cruise control logic is activated when an automated truck is following another vehicle outside of a truck platoon. Following this logic, an automated truck aims to maintain a desired time gap (2 seconds) between itself and the leading vehicle. This logic applies to isolated-automated trucks or connected trucks driving outside a platoon. The car following model is as follows:

\[ a_{Aut}(t) = k_1 [d(t - 1) - t_{des}^{ACC}v(t - 1)] + k_2 [v_{prec}(t - 1) - v(t - 1)] \]  \hspace{1em} (8.2)

Where \( t_{des}^{ACC} \) is the desired time gap in ACC mode (2 sec), \( v_{prec}(t - 1) \) is speed of the preceding vehicle at time \( (t - 1) \), \( k_1 \) and \( k_2 \) are model parameters (0.0561 and 0.3393 respectively).

8.1.3 Cooperative Adaptive Cruise Control (CACC)

Cooperative adaptive cruise control is activated when a connected-automated truck is following another connected-automated truck in a platoon. With this control, the truck maintains a desired time gap in a platoon (1.5 seconds), which is shorter than the desired time gap of adaptive cruise control. Note that if a connected truck is outside a platoon (following a non-connected truck), the connected truck would then follow the ACC logic described above. The CACC car following formula is as follows:

\[ a_{Aut}(t) = k_p e(t - 1) + k_d \dot{e}(t - 1) \]  \hspace{1em} (8.3)
Where $e(t - 1)$ is a measure of deviation from the CACC desired time gap, $t_{\text{des}}^{\text{CACC}}$ is the desired time gap for CACC (1.5 seconds), $\dot{e}(t - 1)$ is derivative of $e(t - 1)$, and $k_p$ and $k_d$ are model parameters (0.0074 and 0.0805 respectively).

8.1.4 Truck Platoon Formation

An opportunistic platoon formation strategy is considered in this case study. An opportunistic formation refers to connected trucks forming platoons whenever possible without inducing any intervention such as pushing certain trucks to change a lane to form a platoon or using reserved lanes. Through this strategy, truck platooning behavior is activated whenever a connected truck is following another connected truck.

8.2 Results and Analysis

This section presents the simulation results of multiple operational scenarios of truck platoons in mixed traffic conditions. The scenarios test the impact of the strategy in low and highly automated traffic conditions. Furthermore, the scenarios analyze the truck platoon sizes formed (number of truck strings) and their duration at different market penetrations of CAVs.

8.2.1 Impact of Automated Truck Platooning in Mixed Traffic Scenarios – Low Automation (30 percent AV) Condition

The scenarios summarized in TABLE 13 evaluate the impact of automated truck platooning at low AV market penetration (30 percent). The scenarios test platooning for two truck
percentages of total traffic: 10 percent and 20 percent. The rest is a mix of AVs and human-driven cars.

**TABLE 13 Automated truck platooning in mixed traffic scenarios – low (30 percent) AV market penetration condition (percent)**

<table>
<thead>
<tr>
<th>Scenario Description</th>
<th>Isolated-Manual Car</th>
<th>Isolated-Automated Car</th>
<th>Isolated-Automated Truck</th>
<th>Connected-Automated Truck</th>
</tr>
</thead>
<tbody>
<tr>
<td>10% Trucks – No Platooning</td>
<td>60</td>
<td>30</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>10% Trucks – Active Platooning</td>
<td>60</td>
<td>30</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>20% Trucks – No Platooning</td>
<td>50</td>
<td>30</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>20% Trucks – Active Platooning</td>
<td>50</td>
<td>30</td>
<td>0</td>
<td>20</td>
</tr>
</tbody>
</table>

FIGURE 44 shows the fundamental diagrams of truck platooning scenarios in low AV market penetration conditions (30 percent). The diagrams show that in both the 10 percent and 20 percent truck percent cases, truck platooning can lead to improvements in traffic throughput (the right two diagrams). This can be caused by the homogenous and less aggressive driving behavior of trucks in platoons. This can also be due to higher traffic density as connected trucks follow each other at shorter distances. The impact of truck platooning on overall travel time, however, is insignificant, as seen in FIGURE 45. Overall travel time refers to the travel time distribution of all vehicles in traffic stream (trucks and cars). This is likely due to the small number of trucks in traffic stream and the opportunistic platoon formation strategy as discussed in the methodology section above (a truck activates platooning whenever it follows another connected truck). Looking at the travel time distribution of trucks only, FIGURE 46, results show that truck platooning can lead to slightly higher travel times for trucks. This can be a result of the less aggressive driving within the truck platoons. On the other hand, truck platooning has no significant impact on the travel time distribution for cars (the majority of traffic stream) as shown in FIGURE 47. The lack of change in the overall travel time distribution indicates insignificant changes in traffic speed,
which confirms that the small improvement in throughput is due to greater traffic density (trucks following each other at smaller distances).

FIGURE 44 Fundamental diagrams for truck platooning scenarios at a low (30 percent) AV market penetration rate.
FIGURE 45 Overall travel time distribution for truck platooning scenarios at a low (30 percent) AV market penetration rate.
FIGURE 46 Truck travel time distribution for truck platooning scenarios at a low (30 percent) AV market penetration rate.

a) 10 percent trucks – no platooning.

b) 10 percent trucks – active platooning.

c) 20 percent trucks – no platooning.

d) 20 percent trucks – active platooning.
FIGURE 47 Car travel time distribution for truck platooning scenarios at a low (30 percent) AV market penetration rate
8.2.2 Impact of Automated Truck Platooning in Mixed Traffic Scenarios – High Traffic Automation (70 percent AV)

The scenarios summarized in TABLE 14 evaluate the impact of automated truck platooning at high AV market penetration (70 percent). The scenarios test platooning for two truck percentages of total traffic: 10 percent and 20 percent. The rest is a mix of AVs and human-driven cars.

<table>
<thead>
<tr>
<th>Scenario Description</th>
<th>Isolated-Manual Car</th>
<th>Isolated-Automated Car</th>
<th>Isolated-Automated Truck</th>
<th>Connected-Automated Truck</th>
</tr>
</thead>
<tbody>
<tr>
<td>10% Trucks – No Platooning</td>
<td>20</td>
<td>70</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>10% Trucks – Active Platooning</td>
<td>20</td>
<td>70</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>20% Trucks – No Platooning</td>
<td>10</td>
<td>70</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>20% Trucks – Active Platooning</td>
<td>10</td>
<td>70</td>
<td>0</td>
<td>20</td>
</tr>
</tbody>
</table>

FIGURE 48 shows the fundamental diagrams for truck platooning scenarios in the high AV market penetration condition (70 percent). Similar to the low automation case, the diagrams show that activating truck platooning can lead to higher traffic throughput. This is because platooned vehicles do not drive aggressively and because they maintain a shorter time gap. FIGURE 49 shows the travel time distributions for the same scenarios. The plots indicate that truck platooning has an insignificant impact on the overall travel time of individual vehicles (no significant shifts in distributions). The small number of trucks in the traffic stream (< 20 percent) is likely the reason for this insignificant change in overall travel time. Looking more deeply into the trucks-only travel time distributions in FIGURE 50, the plots show that platooning could possibly lead to greater travel time for trucks. This is likely due to the less aggressive driving
behavior of platooned vehicles, which also maintaining shorter time gaps between each other compared to non-automated trucks. FIGURE 51 illustrates the cars-only travel time distribution. The plots show that platooning has insignificant impact on travel time for cars. This indicates that the changes in truck driving behavior is not significant enough to impact the car driving behavior and their travel time as the number of trucks is small in the traffic stream (< 20 percent). The insignificant change in travel time (and speed) also indicates that the small increase in traffic throughput is due to higher density (trucks within platoons driving at shorter distances) and not speed itself. While the general trends in the case of high traffic automation is similar to the low traffic automation case discussed previously, the overall throughput is higher and travel time is lower in the case of high automation due to the larger market penetration of AVs.
FIGURE 48 Fundamental diagrams for truck platooning scenarios at a high (70 percent) AV market penetration
FIGURE 49 Overall travel time distribution for truck platooning scenarios at a high (70 percent) AV market penetration rate
FIGURE 50 Truck travel time distribution for truck platooning scenarios at a high (70 percent) AV market penetration rate
FIGURE 51 Car travel time distribution for truck platooning scenarios at a high (70 percent) AV market penetration rate

a) 10 percent trucks — no platooning.

b) 10 percent trucks — active platooning.

c) 20 percent trucks — no platooning.

d) 20 percent trucks — active platooning.
8.2.3 **Truck Platoon Size Analysis**

FIGURE 52 shows the platoon size distribution when trucks make up 10 percent and 20 percent of traffic under both the low and high market penetration rates for automated vehicles. Plots (a) and (b) show that, when 10 percent of traffic is made up of trucks, the platoon size is 2-3 vehicles, while the majority of connected trucks (90 percent) are not in active platoons. Part of the reason for the small sizes is due to the opportunistic nature of platoon formation in this case study where connected trucks activate platooning behavior only if they follow other connected trucks (i.e., platoons are not predefined). The other part of the reason is due to the small number of trucks in the traffic stream, which makes it unlikely that a connected truck would be following another connected truck. Plots (c) and (d) of FIGURE 52 show platoon size when 20 percent of traffic is made up of trucks. In those cases, trucks form platoons more often than in the 10 percent composition scenario, and the range is 2-4 vehicles. This is due to the larger number of trucks in the traffic stream and, therefore, the higher likelihood of platoon formation.
FIGURE 52 Truck platoon size for 10 percent/20 percent trucks at low (30 percent) and high (70 percent) AV market penetration levels
8.2.4 **Truck Platoon Duration Analysis**

FIGURE 53 shows the platoon duration distribution when trucks make up 10 percent and 20 percent of traffic under both the low and high market penetration rates for automated vehicles. Duration in this context refers the numbers of seconds for which a platoon is moving in the traffic stream before it breaks (i.e., a truck leaves a platoon). The plots show that, generally, most platoons break after 50 seconds (220 second is the duration required to cross the study segment at free flow speed). This is due to the small number of trucks in the traffic stream (< 20 percent), the opportunistic nature of platoon formation, and the relatively short travel distance of this urban corridor. Plots 43c and 43d show that platoon duration is slightly longer when trucks make up 20 percent of traffic due to the larger number of trucks in the stream, but still, duration largely remains less than 50 seconds.
Chapter Summary

This chapter presented an analysis of the impact of truck platooning in a mixed traffic environment as an application of a decentralized CAV traffic control strategy. Connected and automated truck platooning simulations show that active platooning can lead to higher traffic throughput due to trucks driving at shorter distances (i.e., headway) in platoons. Platooning, however, seems to have an insignificant impact on overall travel time. The truck platoons formed under the assumed opportunistic platoon formation strategy are of small size (2-4 vehicles) and short duration (mostly less than 50 sec). Under the opportunistic strategy, connected trucks activate

FIGURE 53 Truck platoon duration for 10 percent/20 percent trucks at low (30 percent) and high (70 percent) AV market penetration levels

8.3 Chapter Summary

This chapter presented an analysis of the impact of truck platooning in a mixed traffic environment as an application of a decentralized CAV traffic control strategy. Connected and automated truck platooning simulations show that active platooning can lead to higher traffic throughput due to trucks driving at shorter distances (i.e., headway) in platoons. Platooning, however, seems to have an insignificant impact on overall travel time. The truck platoons formed under the assumed opportunistic platoon formation strategy are of small size (2-4 vehicles) and short duration (mostly less than 50 sec). Under the opportunistic strategy, connected trucks activate
platooning behavior only if they are following other connected trucks. Due to the generally small number of trucks on highways (< 20 percent), forming platoons under this strategy could be difficult, especially over short distances.
9. CONCLUSION AND FUTURE RESEARCH

9.1 Conclusion

The objective of this dissertation is to develop innovative traffic management strategies that utilize the big stream of data generated by CAV systems and the predictive capability of machine learning algorithms. After an overview of the strategic and operational impacts of CAV technology, the dissertation introduced a methodological framework for developing predictive traffic management and control strategies utilizing CAV systems. The framework consists of three main components: 1) traffic monitoring, 2) traffic state prediction, and 3) control strategy. The traffic monitoring component describes how the detailed vehicle trajectories broadcasted by CAVs can be used to estimate traffic properties and track traffic shockwaves (transitions in traffic state) without relying on road sensors. The traffic state prediction component describes how the traffic properties estimated through CAVs can be used to predict future traffic states, a key element in traffic control. Finally, the control strategy component describes two ways in which control actions (e.g. new speed limit, mandatory lane change) can be executed through CAVs.

As an application of the first component of the aforementioned framework, the dissertation presented a novel method to identify traffic shockwave formation and track its propagation based on the speed distribution of individual vehicles available through connected vehicles technology. The method is based on previous theoretical and empirical findings connecting the speed distribution to the onset of flow breakdown, though not specifically related to shock wave formation. The main advantages of this method are: (1) shockwaves can be tracked accurately over small sections of a freeway (e.g., 200-ft segments were used in this analysis), and (2) shockwaves
can be detected clearly for partially connected traffic streams. To test this method, vehicle trajectories from the NGSIM program were analyzed using the US 101 data set for a 2100-ft study segment in Los Angeles, CA. Two main cases were evaluated in this study: full connectivity and partial connectivity. In the case of full connectivity, it was assumed that all vehicles are able to transmit their detailed trajectories. In the partial connectivity case, only a percentage of vehicles were able to transmit their detailed trajectories and shockwaves were tracked using those trajectories only.

Results showed a consistent pattern where shockwave formation (significant reduction in speed propagating over space and time) is associated with a sharp increase in the speed standard deviation value (SSD) that usually occurs before the start of the shockwave development. Results also showed that SSD waves propagates in conjunction with speed shockwaves at almost the same speed, confirming that SSD waves can serve as an indicator to shockwave formation and propagation. These patterns were evident in both the full and partial connectivity cases. In terms of quality of the SSD estimate, the analysis showed that the quality improves at higher market penetrations as the number of vehicles (observations) used to calculate the SSD is higher. Compared to the wavelet transformation method, the analysis showed that SSD waves are more responsive to the reductions in speed than speed wavelet transformations, which can result in a higher shockwave detection accuracy. The higher accuracy in detecting shockwaves in terms of time and space can help improve the effectiveness of active traffic management techniques.

Building on the introduced shockwave detection method, this dissertation presented short-term traffic congestion prediction as an application of the traffic state prediction component in the
general framework for developing predictive CAV traffic management strategies. Two types of predictive models were developed: (1) offline models that are built using historical data only, and (2) online models that are updated in real-time. To build these models, three machine learning techniques were tested: logistic regression, random forests, and neural networks. The analysis of these models showed that the overall prediction accuracy of these models can reach 93%. Furthermore, the results show that congestion can be predicted accurately in the case of partially connected traffic streams, which is important in practice as traffic is unlikely to be fully connected, at least in the early deployment of this new technology.

The congestion prediction models have various safety and traffic performance applications. For instance, the models can be used to warn drivers ahead of traffic slowdowns to prevent potential accidents. In terms of traffic operations, the models can be integrated into traffic control algorithms to enhance their performance. The prediction horizons of these models can be very short for some other applications, given that the prediction accuracy degrades for longer horizons using the current structure of the models. However, the analysis could be extended to predict longer horizons and multiple traffic states given vehicle trajectory data that covers a long duration and a wide range of states, though this matter would need to be empirically tested. Furthermore, as is the case with “black box” statistical methods that are not derived from fundamental theory, the ability to predict changes and states that depart from what had been observed during the training period degrades rapidly. Nonetheless, the high degree of predictability achieved for very short-term forecasts suggests that a productive approach would be to combine the advantages of machine learning techniques with those of approaches based on fundamental traffic concepts and theories.
Utilizing the early shockwave detection method and the congestion prediction models mentioned earlier, the dissertation presented a predictive speed harmonization system with two CAV control strategies: centralized and decentralized. The centralized system relies on a traffic management center to collect information from CAVs within a road segment of interest, predict traffic congestion location, and broadcast updated speed limits to CAVs upstream of that location to mitigate congestion. The system consists of three main modules: 1) Traffic Monitoring, 2) Congestion Prediction, and 3) Speed Control. The Traffic Monitoring module collects detailed traffic trajectories from CAVs and utilizes the early shockwave detection method that tracks changes in speed distribution of CAVs. The Congestion Prediction module identifies the location of the predicted traffic congestion anywhere on the targeted freeway segment. Finally, the Speed Control module determines optimal speed limits to mitigate congestion based on prevailing traffic conditions and broadcasts those limits directly to CAVs upstream of congestion location.

Analysis of multiple operational scenarios showed that activating the predictive speed harmonization system reduces the severity and length of traffic shockwaves and improves the overall stability of traffic. It also improves the overall traffic speed and reduces the average travel time per vehicle. The system performance analysis of partially connected traffic stream indicates that the system’s performance improves at higher connectivity levels. This improvement is due to the higher number of CAV trajectories collected by the system and the more accurate congestion prediction at higher connectivity levels. The system effectiveness analysis in mixed traffic conditions showed that the speed harmonization system improves traffic performance at low market penetration of AVs. At medium to high AV market penetrations, however, the robotic
driving behavior of AVs dominates the traffic stream; therefore, it significantly stabilizes traffic and entirely prevents breakdown.

The decentralized speed harmonization system relies on individual CAVs to collect data through communicating with each other, predict traffic congestion using vehicle-specific models, and adjust their speeds in order to mitigate congestion. The decentralized system consists of three main modules embedded in individual CAVs: 1) Traffic Monitoring, 2) Congestion Prediction, and 3) Speed Control. The Traffic Monitoring module collects detailed traffic trajectories shared by other CAVs within the vehicle’s communication range and utilizes the early shockwave detection method, discussed in chapter 1, to track changes in speed distribution of downstream CAVs. The Congestion Prediction module identifies the location of the predicted traffic congestion downstream of the target vehicle anywhere on the targeted freeway segment. The Speed Control module determines optimal vehicle speed based on downstream traffic conditions.

Results showed that activating the decentralized system reduces the severity of traffic shockwaves, improves stability of traffic, increases overall traffic speed, and reduces travel time. Unlike the centralized system, the decentralized version generated minor shockwaves due to the non-uniformity in predicting traffic congestion and applying control among individual vehicles. However, the overall traffic performance improved for both control strategies. The analysis also showed that having multiple prediction models (fleet-based models) reduces the effectiveness of the strategy. Therefore, a successful application of the decentralized system requires standardization of data collection among vehicles and the ability to communicate with vehicles from other fleets to improve prediction range and accuracy.
Finally, the dissertation presented an opportunistic truck platooning strategy as a special application of decentralized CAV traffic control strategies. Connected and automated truck platooning simulations show that active platooning can lead to higher traffic throughput due to trucks driving at shorter distances (i.e., headway) in platoons. However, platooning as modeled in this work seems to have a limited impact on overall travel time. The truck platoons formed under the assumed opportunistic platoon formation strategy are of small size (2-4 vehicles) and short duration (mostly less than 50 sec). Under the opportunistic strategy, connected trucks activate platooning behavior only if they are following other connected trucks. Due to the generally small number of trucks on highways (< 20 percent), forming platoons under this strategy could be difficult, especially over short distances. Investigating the impact of more concerted platooning approaches, e.g. with pre-formed, longer platoons over longer distances, remains a topic for further work.

9.2 Future Research

Three main factors acting jointly or separately trigger traffic breakdown: 1) high traffic load, 2) bottlenecks, and 3) disturbances caused by individual vehicles (89; 90). As discussed earlier in this dissertation, high traffic loads occur when traffic demand exceeds the sustainable throughput of a road section. A typical example is rush hour congestion. Capacity reductions or “bottlenecks” may be permanent, such as on-ramps and off-ramps, or temporary such as traffic accidents or slow-moving vehicles. As for traffic disturbances, those refer to temporary perturbations in the traffic flow. The predictive speed harmonization strategies presented in this dissertation focus on the first factor of the three. Once traffic congestion is predicted, the strategies
reduce the speeds of CAVs approaching the congestion location and therefore they reduce the overall traffic load at that location, temporarily until congestion is resolved.

Another approach to reducing the likelihood of a traffic breakdown that can be developed in future work would focus on minimizing the disturbances caused by individual vehicles. Traffic perturbations can be caused by lane-change maneuvers, abrupt braking, speeding, or long-lasting overtaking maneuvers of trucks. While such actions are difficult to predict for traffic control purposes, they often create speed variations that can be estimated from CAV-generated data, as shown in Chapter 4, and potentially be predicted.

Similar to the presented speed harmonization strategies, a traffic disturbance minimization strategy would utilize the information broadcasted by CAVs to estimate the speed standard deviation of traffic in real-time. Instead of predicting traffic congestion, however, the new strategy would predict future traffic disturbances for all road segments by predicting its proxy, the traffic speed standard deviation. Once disturbances are predicted to occur, the strategy would identify the vehicles that are likely to cause those traffic disturbances by, for example, measuring their speeds and acceleration relative to other vehicles (aggressiveness), and send out advisory speed to those vehicles in order to prevent or minimize potential disturbances.

The traffic disturbance minimization strategy can potentially be developed for centralized and decentralized systems. A centralized system would monitor traffic perturbations in a road segment of interest by collecting and analyzing information broadcasted by all CAVs within that segment. A decentralized system, on the other hand, relies on individual vehicles to predict traffic perturbations ahead of them or within their vicinity and adjust their speeds to minimize them.
Traffic disturbance minimization can be considered a preventive strategy that can be coupled with a speed harmonization strategy. In this case, the disturbance minimization strategy would delay breakdown as much as physically possible while the speed harmonization strategy would accelerate traffic recovery once congestion is realized.

In addition to developing new CAV control strategies, such as traffic disturbances minimization, future work can also focus on improving the speed harmonization systems developed in this dissertation. One particular area that can be improved is the traffic state prediction component of the speed harmonization system. The current prediction models assume a binary traffic state, congested and uncongested. By predicting more traffic states, such as Kerner’s three traffic phases (141), and applying different control strategies to each of those phases, the overall traffic performance can potentially be improved. The developed strategies can also be improved by expanding their control actions beyond changing vehicle speeds such as changing lanes or routes.

Another area to explore in future work is integrating the developed speed harmonization strategies with other active traffic control strategies such as queue or collision warning. Studies have shown that implementing a combination of active strategies, such as speed harmonization and ramp metering can outperform the implementation of individual strategies (61; 65).
REFERENCES


