

University of Central Florida

**STARS**

---

Electronic Theses and Dissertations, 2020-

---

2023

## Optimizing Information Values in Smart Mobility

Fatima Afifah

*University of Central Florida*



Part of the [Civil Engineering Commons](#)

Find similar works at: <https://stars.library.ucf.edu/etd2020>

University of Central Florida Libraries <http://library.ucf.edu>

This Doctoral Dissertation (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations, 2020- by an authorized administrator of STARS. For more information, please contact [STARS@ucf.edu](mailto:STARS@ucf.edu).

---

### STARS Citation

Afifah, Fatima, "Optimizing Information Values in Smart Mobility" (2023). *Electronic Theses and Dissertations, 2020-*. 1872.

<https://stars.library.ucf.edu/etd2020/1872>

# OPTIMIZING INFORMATION VALUES IN SMART MOBILITY

by

FATIMA AFIFAH

B.S. Bangladesh University of Engineering and Technology, 2016

M.S. Bangladesh University of Engineering and Technology, 2019

A dissertation submitted in partial fulfilment of the requirements  
for the degree of Doctor of Philosophy  
in the Department of Civil, Environmental and Construction Engineering  
in the College of Engineering and Computer Science  
at the University of Central Florida  
Orlando, Florida

Summer Term  
2023

Major Professor: Zhaomiao Guo

© 2023 Fatima Afifah

## ABSTRACT

Smart mobility, enabled by advanced sensing, communication, vehicle, and emerging mobility technologies, has transformed transportation systems. Real-time information shared by public and private entities plays a pivotal role in smart mobility, which facilitates informed decision-making, including effective mode choice, dynamic vehicle control, optimized travel routing, and strategic vehicle relocation. While more information is believed to benefit individual decision makers, it is crucial to acknowledge that the effects of information on transportation network performance are contingent; more information may not always benefit the safety and mobility of the whole system.

The goal of this dissertation is to investigate the effects of information shared by public and private transportation entities on system-level performance. The challenges are primarily due to the lack of a unified modeling framework to endogenously reflect the decentralized multi-agent interaction involved in the interconnected transportation networks and the resulting computational complexities arising from non-convexity and high dimensionality. To address these challenges, this dissertation proposes novel modeling frameworks and computational solutions for three cutting-edge smart mobility applications. First, to examine the impact of en-route information on a transportation network, we propose a novel two-stage stochastic traffic equilibrium model to characterize the equilibrium traffic patterns considering adaptive routing behavior when locational en-route traffic information is provided through infrastructure-to-vehicles (I2V) communications. This model is formulated as a convex stochastic optimization problem so that efficient stochastic programming algorithms can be directly leveraged to achieve scalability. Second, to achieve optimal control over real-time variable speed limits information sharing and evaluate its impact on the network, we propose a twin-delayed deep deterministic policy gradient model, which converges more reliably than state-of-the-art deep reinforcement learning models. We investigate the transferability of the control algorithm and conduct comparative analyses of different traffic control strategies and spatial

distributions of variable speed limit control (VSLC) deployment Third, to assess the impacts of information provided by private ride-sourcing companies on transportation network congestion, we propose a Stackelberg framework for spatial pricing of ride-sourcing services considering traffic congestion and convex reformulation strategies under mild conditions. We perform numerical experiments on transportation networks of varying scales and with diverse transportation network company (TNC) objectives, aiming to derive policy insights regarding the implications of spatial pricing information on transportation systems.

## **ACKNOWLEDGMENTS**

I would like to express my deepest appreciation and gratitude to my honorable supervisor, Dr. Zhaomiao Guo, for his invaluable guidance and unwavering support throughout the duration of this dissertation. His exceptional supervision, insightful advice, and profound wisdom have been instrumental in shaping and completing this work in a timely manner. I am truly fortunate to have had the privilege of working under his mentorship.

I would also like to extend my heartfelt thanks to my family for their constant encouragement and unwavering belief in my abilities. Their love and support have been a constant source of strength throughout this journey.

Additionally, I am deeply grateful to my best friend, Amanda Scott, whose steadfast support, encouragement, and understanding have been a source of motivation and inspiration. Her presence and positive energy have been a source of solace during challenging times.

# TABLE OF CONTENTS

LIST OF FIGURES . . . . .	xiii
LIST OF TABLES . . . . .	xvi
CHAPTER 1: INTRODUCTION . . . . .	1
1.1 Background . . . . .	1
1.2 Motivation . . . . .	2
1.3 Challenges . . . . .	7
1.4 Dissertation Objectives . . . . .	7
1.5 Contributions . . . . .	8
1.6 Structure of the Dissertation . . . . .	9
CHAPTER 2: SYSTEM-LEVEL IMPACTS OF EN-ROUTE INFORMATION SHARING CONSIDERING ADAPTIVE ROUTING . . . . .	12
2.1 Introduction . . . . .	12
2.2 Literature Review . . . . .	13
2.2.1 Safety and Mobility Implications of CAVs . . . . .	14
2.2.2 Adaptive Decision Making with Information Updates . . . . .	14

2.3	Methodology . . . . .	16
2.3.1	Network Modeling . . . . .	17
2.3.1.1	Beckmann’s Models for UE . . . . .	17
2.3.1.2	Adaptive Traffic Assignment Models . . . . .	18
2.3.1.3	Illustrative Example for Two-stage SUER . . . . .	23
2.3.2	Parametric Estimation . . . . .	26
2.3.2.1	Link Performance Function . . . . .	27
2.3.2.2	Collision Risk Function . . . . .	28
2.3.3	Network Safety Analyses . . . . .	31
2.3.3.1	Representation of Information Sharing Locations . . . . .	31
2.3.3.2	Network Safety Assessment . . . . .	32
2.4	Numerical Results . . . . .	32
2.4.1	Base Case Results . . . . .	33
2.4.2	Sensitivity Analyses . . . . .	37
2.4.2.1	Sensitivity of Information Sharing Strategies . . . . .	37
2.4.2.2	Sensitivity of O-D Demand . . . . .	42
2.4.2.3	Sensitivity Analysis on Incident Severity . . . . .	43



2.5	Discussion . . . . .	44
CHAPTER 3: OPTIMAL SPEED LIMIT CONTROL FOR NETWORK MOBILITY AND SAFETY: A TWIN-DELAYED DEEP DETERMINISTIC POLICY GRADIENT APPROACH . . . . .		
		46
3.1	Introduction . . . . .	46
3.2	Literature Review . . . . .	47
3.3	Problem Statement . . . . .	51
3.4	Theoretical Background . . . . .	52
3.4.1	Basic Overview of Reinforcement Learning . . . . .	53
3.4.2	Deep Reinforcement Learning . . . . .	58
3.4.2.1	Experience Replay . . . . .	58
3.4.2.2	Target network . . . . .	59
3.4.2.3	Deep Deterministic Policy Gradient . . . . .	61
3.5	Proposed Algorithm . . . . .	63
3.5.1	Deep Network Architecture . . . . .	64
3.5.2	Online Update mechanism . . . . .	66
3.5.2.1	Online Critic Network Update . . . . .	66
3.5.2.2	Online Actor Network Update . . . . .	67

3.5.3	Target Update mechanism . . . . .	68
3.5.4	Exploration method . . . . .	68
3.5.5	State, Reward, and Action for the TD3 model . . . . .	69
3.5.5.1	State: . . . . .	70
3.5.5.2	Reward: . . . . .	70
3.5.5.3	Action: . . . . .	74
3.5.6	TD3 configuration and algorithm summary . . . . .	74
3.5.7	SUMO configuration for the TD3 model . . . . .	75
3.6	Experiment Results . . . . .	77
3.6.1	Base Case Results . . . . .	78
3.6.2	Comparison with other Methods . . . . .	81
3.6.3	Assessment of Different Traffic Control Strategies in the Network . . . . .	83
3.6.4	Assessment of Generalization of the Agent . . . . .	88
3.6.5	Sensitivity of VSLC placement in the Network . . . . .	92
3.7	Discussion . . . . .	105
CHAPTER 4: SPATIAL PRICING OF RIDE-SOURCING SERVICES IN A CONGESTED TRANSPORTATION NETWORK . . . . .		106

4.1	Introduction . . . . .	106
4.2	Literature Review . . . . .	107
4.3	Methodologies . . . . .	110
4.3.1	Mathematical Modeling: Ignoring Waiting Time . . . . .	111
4.3.1.1	Traffic Modeling . . . . .	111
4.3.1.2	TNC Behaviors . . . . .	117
4.3.2	Mathematical Modeling: Considering Waiting Time . . . . .	119
4.4	Solution Approach . . . . .	125
4.4.1	Solution Approach: Ignoring Waiting Time . . . . .	125
4.4.2	Solution Approach: Considering Waiting Time . . . . .	128
4.5	Numerical Examples . . . . .	131
4.5.1	Three-node Network . . . . .	132
4.5.1.1	Ignoring Waiting Time . . . . .	133
4.5.1.2	Considering Waiting Time . . . . .	136
4.5.2	Sioux Falls Network . . . . .	138
4.5.2.1	Ignoring Waiting Time . . . . .	138
4.5.2.2	Considering Waiting Time . . . . .	140

4.6	Discussion . . . . .	141
CHAPTER 5: CONCLUSIONS . . . . .		144
5.1	Summary of Major Results . . . . .	145
5.2	Future Research Direction . . . . .	146
APPENDIX A: PROOF OF SYSTEM-LEVEL IMPACTS OF EN-ROUTE INFORMATION SHARING CONSIDERING ADAPTIVE ROUTING . . . . .		
		149
A.1	Proof of Theorem 2.3.1 . . . . .	150
A.1.1	Sufficient Condition of Theorem 2.3.1 . . . . .	151
A.1.2	Necessary Condition of Theorem 2.3.1 . . . . .	153
APPENDIX B: PROOF , EXPLANATION AND EXTENSION OF SPATIAL PRICING OF RIDE-SOURCING SERVICES IN CONGESTED TRANSPORTATION NETWORK . . . . .		
		156
B.1	Proofs. . . . .	157
B.2	Relationship between Waiting Time and Driver/Rider Flow . . . . .	162
B.3	Extensions . . . . .	163
B.3.1	Extension of Linear Demand Function . . . . .	163
B.3.2	Extension of Network Effects of Available Drivers . . . . .	166

REFERENCES . . . . . 190

## LIST OF FIGURES

2.1	Four-node Network with Stochastic Link Cost . . . . .	23
2.2	Convergence of Traffic Patterns and Expected Travel Time . . . . .	26
2.3	Study Area in Orlando . . . . .	34
2.4	Link Flow Changes Between Normal Scenario and Incident Scenario for the Base Case . . . . .	35
2.5	Relative Difference of Congestion Level between Information Sharing and No Information Sharing . . . . .	36
2.6	Comparison of Collision Risk at each Link between Information Sharing vs. No information Sharing . . . . .	37
2.7	Effect of information sharing at individual nodes separately . . . . .	39
2.8	Impact of Perfect and no Information Sharing . . . . .	40
2.9	Impact of at Information Sharing at Selected Nodes . . . . .	41
2.10	Effect of Change in O-D Demand . . . . .	43
2.11	Effect of Incident Severity on Network Mobility and Safety . . . . .	44
3.1	Problem Setting . . . . .	52
3.2	Traffic oscillation in the routing links . . . . .	53

3.3	Interaction between agent and environment . . . . .	54
3.4	Structure of TD3 . . . . .	65
3.5	Logarithmic function . . . . .	72
3.6	Base Case Results . . . . .	79
3.7	Speed Limit for the Base Case . . . . .	80
3.8	Traffic oscillation when controlled . . . . .	86
3.9	Traffic oscillation when controlled . . . . .	87
3.10	Speed profile of Link 2-3 for different traffic control strategies . . . . .	87
3.11	Effect of Different Traffic Attributes on Network Safety and Mobility . . . . .	90
3.12	Position of VSL for improving network safety and mobility . . . . .	94
3.13	Performance of different VSLLC placements to improve network safety . . . . .	95
3.14	Traffic behaviors for different VSLLC placements for Network Safety Improvement . . . . .	97
3.15	Speed profile for different VSLLC placements to improve network safety . . . . .	99
3.16	Performance of different VSLLC placements to improve network mobility . . . . .	101
3.17	Traffic behaviors for different VSLLC placements for Network Mobility Improvement . . . . .	103
3.18	Speed profile of different VSLLC implementation for improving mobility . . . . .	104

4.1	Illustration of Network Augmentation . . . . .	123
4.2	Three Nodes Test Network . . . . .	132
4.3	Impact of Ride-sourcing Service on Traffic Congestion . . . . .	134
4.4	Impacts of Supply and Demand on Optimal Surge Pricing . . . . .	136
4.5	Convergence Patterns with Different Starting Points $\rho^0$ . . . . .	137
4.6	Equilibrium Solutions For Ignoring and Considering Waiting Time Cases . . . . .	138
4.7	Sioux Falls Test Network . . . . .	139
4.8	Impacts of Supply and Demand on Optimal Surge Pricing . . . . .	140
4.9	Convergence of ES and $\rho$ Considering Waiting Time . . . . .	141
B.1	Graphical Representation of Waiting Time . . . . .	162
B.2	Convergence of Available Drivers Distribution . . . . .	167



## LIST OF TABLES

2.1	Equilibrium solutions with risk-neutral adaptive behaviors . . . . .	25
3.1	TD3 Parameters . . . . .	75
3.2	Average Performance of Different Models in Improving Network Safety and Mobility in Base Case . . . . .	82
3.3	Average Performance of Different Traffic Control Strategies . . . . .	84
3.4	Average Performance of Different VSLC Placement for Network Safety Im- provement . . . . .	96
3.5	Average Performance of Different VSLC Placement for Network Mobility Improvement . . . . .	102
B.1	Sample Regression Input . . . . .	163
B.2	Regression Summary Output of (4.8) . . . . .	164

# CHAPTER 1: INTRODUCTION

## 1.1 Background

Smart mobility, driven by advanced sensing, communication, computational, and emerging mobility technologies, offers transformative solutions to address the challenges of mobility and safety in transportation systems. Congestion on roadways and traffic accidents incur significant economic, environmental, and social costs. According to the 2021 Urban Mobility Report by the Texas A&M Transportation Institute, the average commuter experienced 54 hours of traffic delay in 2019, costing \$210 billion (TTI, 2021). This congestion also resulted in the wastage of 3.5 billion gallons of fuel, contributing to air pollution and climate change. The National Highway Traffic Safety Administration (NHTSA) reported that in the same year, there were 36,500 fatalities, 4.5 million injuries, and 23 million damaged vehicles, amounting to a cost of \$340 billion (NHTSA, 2023). In today's information-rich era, where a wealth of data is readily available, the optimization of information values in smart mobility holds immense promise for improving the safety and efficiency of transportation networks. This dissertation investigates the impact of information shared by both public and private entities and optimizes its benefits to enhance the safety and efficiency of transportation networks.

Smart mobility is a paradigm within the transportation sector that addresses the challenges of congestion and accidents by harnessing innovative solutions based on information and communication technologies (ICT). Smart mobility encompasses a range of solutions, including connected and automated vehicles (CAVs), infrastructure-to-vehicles (I2Vs), variable speed limit controls (VSLCs), and shared mobility services such as ride-sourcing services. These different aspects of smart mobility work together to create a seamless and interconnected transportation system.

Real-time information sharing plays a crucial role in enhancing smart mobility solutions. It involves collecting, analyzing, and disseminating up-to-date data about traffic conditions, road closures, parking availability, public transit schedules, and other relevant information. This real-time data is made accessible to users through mobile applications and digital displays, empowering them to make informed decisions. By leveraging real-time information, drivers can choose the most efficient routes, avoiding congested areas and reducing travel time. Additionally, transportation management systems can dynamically respond to changing conditions by adjusting speed limits on VSLCs, optimizing the safety and efficiency of the road network. Recently, ride-sourcing services have revolutionized transportation by leveraging real-time information shared through mobile platforms, benefiting both drivers and riders. Transportation network companies (TNCs) provide pricing information in real-time. This dynamic pricing information improves resource allocation and reduces waiting times for riders while enabling drivers to maximize their profitability. Better understanding the impacts of information sharing on smart mobility systems is critical for guiding the design, planning, operation, and policy decision-making processes, ultimately leading to more efficient and sustainable transportation networks.

## **1.2 Motivation**

While information sharing can improve travelers' decision-making and routing behavior (Bonsall, 1992; Chorus et al., 2006), as well as improve road capacity and travel-time savings (Levinson, 2003; Lu et al., 2011; Toledo and Beinhaker, 2006; Tsirimpa et al., 2007), information sharing may not always benefit the network performance. Studies have shown that sharing too much information can lead to the "information paradox" where sharing more information can worsen network performance (Acemoglu et al., 2018; Afifah et al., 2023; Liu and Liu, 2018; Unnikrishnan and Waller, 2009; Wijayaratna et al., 2017). For instance, Unnikrishnan and Waller (2009) in

their study showed that system performance deteriorates when drivers receive information about all the arcs in the network, while Acemoglu et al. (2018) found that the impact of additional information on network performance depends on the network configuration and specific conditions of information provision. Liu and Liu (2018) in their study showed that optimal information sharing provision can reduce the total system cost. Therefore, it is crucial to develop effective modeling frameworks to assess and optimize the effects of information sharing on transportation networks.

In the context of smart mobility, several specific areas have received attention in relation to information sharing. One such area is I2V communication, which involves sharing real-time information with vehicles to improve transportation efficiency and safety. Additionally, the use of VSLC systems has emerged as a means of dynamically adjusting speed limits based on real-time traffic conditions. Another notable area is ride-sourcing services, including TNCs, which have introduced new forms of information sharing such as real-time traffic updates and pricing data for both riders and drivers.

The advent of CAVs has spawned new possibilities for exchanging information between infrastructure, such as traffic signals or roadway sensors, and the vehicles themselves. I2V communication is an information-sharing strategy that allows CAVs to receive real-time information about traffic conditions, road hazards, and other relevant data from surrounding infrastructure and other connected vehicles, enabling them to make informed and optimal decisions. However, while the existing literature primarily focuses on the developmental aspects of I2V technology with limited attention given to the examination of information sharing impact (Pérez et al., 2010; Sukuvaara and Nurmi, 2009), studies on CAVs have demonstrated the potential benefits of implementing CAVs in improving the safety and efficiency of transportation systems (Fagnant and Kockelman, 2015; Liu and Fan, 2021; Ma et al., 2021; Papadoulis et al., 2018; Rahman et al., 2019; Yue et al., 2018; Zhao et al., 2021)). For instance, Rahman et al. (2019) examined the impact of CAVs on the crash risks of arterials, focusing on segment and intersection crashes using surrogate safety assessment

techniques and microsimulation.

It is worth noting that the adaptive capability of CAVs receiving en-route information from I2V technologies has not been extensively explored in the existing studies. Additionally, most of the existing studies consider CAVs to improve safety and mobility at local levels<sup>1</sup>. However, focusing on local-level impacts may cause biases since a transportation system is an interconnected network. Local node or link transportation state changes, such as the occurrence of weather events, road construction, traffic management activities, traffic accidents, etc, will not only impact the link performance functions (used for mobility measures) and crash risk (used for safety measures) for the impacted links but also have broader impacts on other links due to vehicles rerouting and traffic redistribution. For instance, if congestion happens on one link, sharing information on this event leads CAVs to use other alternative paths, thus contributing to congestion on those alternative paths at a broader scale. These gaps in the existing studies motivate us to develop a methodological approach to investigate the impact of I2V technologies on the transportation network considering the rerouting capabilities of CAVs.

Another promising information sharing strategy in the realm of smart mobility is the use of VSLC. It involves dynamically adjusting speed limits in response to traffic conditions and disseminating them to vehicles on the road, which can help regulate traffic flow and reduce congestion. Prior research on speed limit control has primarily focused on the impact of speed limit changes on safety and mobility at the individual road segments level such as a local intersection or link level (Abdel-Aty et al., 2006; Hellinga and Mandelzys, 2011; Islam et al., 2013; Kušić et al., 2020a; Lin et al., 2004; Wang et al., 2019a; Wu et al., 2020c). For example, Wu et al. (Wu et al., 2020c) used a deep deterministic policy gradient algorithm to alleviate recurrent bottlenecks at a freeway section

---

<sup>1</sup>Local-level (i.e., link/node-level) safety and mobility refer to the crash risk and congestion on neighboring road segments or intersections only, while system-level safety and mobility considers the interconnection of the transportation network and refers to the measures for the whole transportation network by aggregating the local-level safety and mobility for each link/node.

and found that VSLC can improve safety and mobility. However, this study does not consider the potential impacts of VSLC on the surrounding links or the effect of VSLC implementation on a network.

Optimizing the implementation of VSLC strategies in a transportation network can be daunting due to the complex and dynamic nature of the network topology and the adaptive decision-making process involved. Identifying the optimal locations for implementing VSLC strategies in a transportation network to maximize both network mobility and safety presents a challenging problem that has yet to be fully explored. Furthermore, the use of VSLC in transportation networks has the potential to affect the routing decisions of drivers by modifying speed limits in response to current traffic conditions. This adjustment can influence the attractiveness of certain routes over others, which may result in drivers switching routes. This can eventually lead to the increased total travel time of the entire system and cause traffic oscillations between alternative routes. In addition, frequent lane changing due to vehicle rerouting can contribute to a higher risk of collisions and local bottlenecks. Despite the potential of VSLC in mitigating congestion and enhancing safety, an in-depth analysis of VSLC implementations on a transportation network and the impact of this dynamic information on traffic behavior remains lacking. This motivates us to evaluate the potential effects of VSLC placement on both network safety and mobility.

The emergence of ride-sourcing services has brought attention to a new type of information sharing, which includes providing real-time traffic updates and sharing data about drivers, vehicles, and pricing to users. Many studies have been conducted to investigate how TNCs determine their pricing strategies. Optimizing pricing strategies for TNCs is a complex task due to the interplay of space and time, as well as the potentially conflicting interests between TNCs and society. While dynamic pricing has been proposed to balance supply and demand, existing research has predominantly focused on the temporal aspect. However, pricing decisions in one location can potentially have ripple effects on the entire transportation network, as demand and supply markets

are interrelated. To address this issue, spatial pricing strategies have been suggested to account for long-term, predictable, and uneven demand patterns (Bimpikis et al., 2019). From a network perspective, studying dynamic pricing is crucial for both TNCs and society's welfare. Nonetheless, while TNCs prioritize profits, their pricing strategies may not necessarily align with societal interests, which include the welfare of riders, drivers, and other transportation users. Therefore, it is essential to better comprehend the effects of TNCs' pricing decisions on the overall transportation system to develop regulatory guidelines and planning strategies for policymakers.

In addition to the challenges of determining optimal pricing strategies for TNCs, ride-sourcing services have raised concerns about their impact on traffic congestion. Studies have found ride-sourcing services have contributed to traffic congestion in several cities. For instance, Qian et al. (2020) used ride-sourcing data to demonstrate that increased demand for these services is a leading cause of traffic congestion in New York City. Similarly, Castiglione et al. (2018) reported a 50% increase in congestion in San Francisco due to ride-sourcing services. To mitigate this issue, several cities have imposed congestion surcharges on ride-sourcing services, including NYC (NYCTLC, 2019) and Chicago (Pierog, 2019). However, while a larger fleet size of ride-sourcing services can reduce waiting time, it can also intensify congestion, as found by Beojone and Geroliminis (2021) using Chinese ride-sourcing data. Furthermore, research has shown that younger people tend to favor ride-sourcing services over owning private vehicles, suggesting that the market share of these services will likely increase in the future (Alemi et al., 2018; Dzisi et al., 2020). The limitations in the current body of literature prompt us to examine the effects of ride-sourcing services on traffic congestion through the development of a comprehensive methodological framework that encompasses key pertinent stakeholders (i.e., drivers, riders, and TNCs), with the aim of informing policy regulation on TNC.

### **1.3 Challenges**

The complexity of interconnected transportation networks and the computational challenges they pose have hampered the development of unified modeling frameworks incorporating decentralized multi-agent interaction to assess the impact of information sharing on system-level performance. The complexity of interconnected transportation networks arises from the interactions among multiple agents and network components, as well as the heterogeneity in network characteristics and traffic conditions. These complexities further pose significant computational challenges; non-convexity and high dimensionality make finding optimal solutions difficult, requiring scalable computational techniques.

To overcome these challenges, this dissertation proposes various methodological frameworks incorporating decentralized multi-agent interaction with information updates and scalable solution approaches to investigate the impact of information sharing by public and private transportation entities on the safety and mobility of transportation networks. This dissertation serves as a fundamental contribution to the field of research on the impact of information in transportation networks within the context of smart mobility. It lays the foundation by developing various modeling frameworks and computational solutions aimed at comprehending the effects of information from various smart mobility applications, notably, I2V, VSLC, and ride-sourcing services, on transportation networks.

### **1.4 Dissertation Objectives**

This dissertation focuses primarily on investigating the impact of information sharing by public and private transportation entities on system-level safety and mobility, with a particular interest to address the modeling and computational challenges associated with assessing the information



shared by diverse state-of-the-art smart mobility applications. The dissertation focuses on the following specific objectives:

- I. Develop a novel and computationally tractable transportation network model that captures the equilibrium traffic patterns while accounting for the adaptive routing of CAVs with en-route information updates provided by I2V technologies, in order to investigate the impact of en-route information on a transportation network.
- II. Develop a deep reinforcement learning model for VSLC implementation, in order to achieve optimal control over the information shared with drivers regarding speed limits and evaluate its impact on a transportation network performance.
- III. Develop a Stackelberg framework for spatial pricing information of ride-sourcing services incorporating the diverse interests of key stakeholders, namely TNC, riders, and drivers while considering traffic congestion, to analyze the impact of information shared by private entities on transportation congestion.

## **1.5 Contributions**

The dissertation has made the following contributions:

- I. This dissertation will provide a better understanding of the impact of en-route information on a transportation network considering adaptive routing behavior, which is lacking in the existing studies. In Chapter 2, we propose a scalable modeling framework to characterize the equilibrium traffic patterns considering adaptive routing behavior with locational en-route traffic information. The model is structured as a convex stochastic optimization problem,

enabling the utilization of efficient stochastic programming algorithms for achieving scalability in the optimization process.

- II. This dissertation distinguishes itself from previous research by addressing the optimal control of VSL from a network-level perspective, considering both mobility enhancement and safety improvement. In Chapter 3, we develop a deep actor-critic framework that allows for the effective control of VSLCs in a network, taking into account the rerouting behavior of vehicles. By incorporating the rerouting behavior of vehicles, the framework considers the dynamic nature of traffic flow and the potential impacts of VSLCs on driver decisions regarding route choices. Additionally, we propose suitable reward functions that enable a comprehensive evaluation of network performance, encompassing mobility, safety, and traffic oscillation.
- III. In Chapter 4, we propose a Stackelberg framework for spatial pricing of ride-sourcing services considering traffic congestion with convex reformulation. By doing so, this dissertation distinguishes itself from existing research in two key aspects. First, we incorporate the relocation of drivers and transportation congestion explicitly, enabling accurate modeling of ride-sourcing patterns and comprehensive evaluation of spatial pricing strategies on congestion. Second, we address the computational challenges associated with optimizing spatial pricing, which arise from the inherent bi-level structure of the problem. To overcome these challenges, we propose effective computational algorithms accompanied by rigorous analyses concerning the existence, uniqueness, and convergence of the optimal solutions.

## **1.6 Structure of the Dissertation**

The remainder of the dissertation is divided into four chapters which shows how each chapter contributes to the goal of understanding the impact of information sharing on transportation net-

work performance. From chapter two to chapter four, each chapter addresses specific objectives in the context of state-of-the-art smart mobility applications, emphasizing the limitations of prior research, proposing methodological frameworks and solution approaches, and conducting evaluations to contribute to the existing literature and practical applications by bridging knowledge gaps and providing valuable insights into the impact of information sharing in smart mobility.

In Chapter 2, we propose a novel two-stage stochastic traffic equilibrium model that captures the equilibrium traffic patterns by incorporating adaptive routing behavior in response to en-route traffic information. The proposed methodology is evaluated using real traffic and accident data from the Orlando transportation network. We analyze diverse I2V information sharing strategies to assess the impacts of different I2V information sharing locations on transportation network safety and mobility. Based on our analysis, we provide insights into the optimal design of traffic information sharing systems. This chapter contributes to the objective of investigating the impact of en-route information sharing on transportation network performance.

In Chapter 3, we propose a deep reinforcement learning (DRL) model to address the gap in understanding the impact of VSLC on network-level performance in transportation systems. Our DRL-based VSLC controller is designed to handle a large number of possible speed limits at each time step using a deep actor-critic framework. Additionally, we propose effective reward functions to characterize network mobility, safety, and traffic oscillation. Using a network we demonstrate the generalizability of the control algorithm across different traffic and driving behavior attributes as well as the performance of the proposed model compared with state-of-the-art DRL models. Furthermore, we assess the efficacy of various control algorithms and analyze the impacts of VSLC implementations in different locations, with a focus on improving overall network mobility and safety. Overall, this chapter contributes to the objective of achieving optimal control over speed limit information shared by VSLC and examining its impact on network mobility and safety.

In Chapter 4, we analyze the effects of spatial pricing for ride-sourcing services in a Stackelberg framework while considering traffic congestion. We employ a combined distribution and assignment approach at the lower level to explicitly model the interactions between drivers' relocation, riders' mode choice, and all travelers' routing decisions. At the upper level, a single TNC determines spatial pricing strategies to minimize imbalance in a two-sided market. Our analysis reveals the existence of optimal pricing strategies for locational imbalance minimization, and we propose effective algorithms with reliable convergence properties. Additionally, we demonstrate that the optimal pricing is unique and can be solved in a convex reformulation when waiting time is small compared to travel time. We conduct numerical experiments on various scales of transportation networks with different TNC objectives to generate policy insights regarding how spatial pricing could affect transportation systems. This chapter contributes to the specific objective of investigating the impact of information shared by private entities on transportation network congestion.

Finally, in Chapter 5, we conclude the dissertation by discussing the overall findings of the chapters and providing potential directions for future studies for information sharing in smart mobility.

## **CHAPTER 2: SYSTEM-LEVEL IMPACTS OF EN-ROUTE INFORMATION SHARING CONSIDERING ADAPTIVE ROUTING**

### **2.1 Introduction**

The increasing market penetration of connected and automated vehicles (CAVs) is expected to reach 8 million by 2025 (ABI, 2018). Although fully automated vehicles are still under development, advancements in vehicle connectivity and onboard computation have the potential to transform transportation systems (Elliott et al., 2019). Infrastructure-to-vehicle (I2V) technology plays a crucial role in providing CAVs with real-time information for making optimal travel decisions (Böhm and Frötscher, 2009). Different strategies for sharing information among vehicles significantly influence individual decision-making and collectively determine network performance. While existing studies have demonstrated the benefits of CAVs at local levels, it is important to consider the interconnected nature of transportation networks. Changes in local transportation conditions can have broader impacts on other links and nodes due to traffic rerouting and redistribution. Research focusing on the system-level safety and mobility implications of CAVs considering an interconnected network is limited. Understanding the effects of en-route information sharing on overall system performance requires an analytical modeling framework that incorporates traffic patterns, information updates, and traffic rerouting. By studying system-level impacts, we can avoid myopic decision-making and recognize that increased information does not always lead to improved safety and mobility. This motivation aligns with classic transportation paradoxes (e.g., Barass's paradox (Rapoport et al., 2009) and Knight-Pigou-Downs paradox (Dechenaux et al., 2014; Morgan et al., 2009)), highlighting the need to consider the broader system effects of local infrastructure improvements.

In this dissertation, we aim to investigate the impacts of locational en-route information updates on transportation system safety and mobility considering CAVs' adaptive routing. Note that we do not require autonomous driving features. Instead, we leverage advanced vehicle information sensing and computational technologies that can assist vehicles to make routing decisions, which can be executed by either human-driven vehicles or automated vehicles. While various types of information can be shared, this study focuses on sharing mobility information on specific transportation links with CAVs when they pass by I2V devices. Both the prior knowledge and real-time information on the transportation states will be considered by CAVs to determine the optimal individual (re)routing decisions, which collectively determine the network traffic equilibrium patterns. The contributions of this study are two-fold.

1. We propose a novel and computationally tractable transportation network model to describe the traffic equilibrium patterns considering the adaptive routing of CAVs with en-route information updates.
2. We evaluate the impacts of information sharing locations on transportation safety and mobility at a network level with real-world traffic data considering adaptive routing of CAVs.

## **2.2 Literature Review**

I2V and CAV technologies provide tremendous opportunities to improve traffic safety and mobility through information sharing on real-time transportation states. In this section, we review relevant literature on the following two aspects: (1) safety and mobility implications of CAVs; and (2) traffic network modeling with the adaptive decision making of CAVs with information updates.

### 2.2.1 Safety and Mobility Implications of CAVs

Extensive studies have been conducted to demonstrate the effectiveness of safety and mobility improvements leveraging different levels of CAVs. Majority of these existing studies are limited to local safety and mobility improvements, such as motorway (Papadoulis et al., 2019), freeway (Amini et al., 2021; Xie et al., 2017; Zhao et al., 2021; Zheng et al., 2019), and signalized/unsignalized intersection (Du et al., 2018; Fyfe and Sayed, 2017; Lee and Park, 2012; Ma et al., 2021; Morando et al., 2018; Rahman et al., 2019; Wang et al., 2019c). For example, Papadoulis et al. (2019) developed a CAV control algorithm and simulated it in VISSIM for motorway segment safety evaluation. Zheng et al. (2019) demonstrated freeway safety and mobility improvement by proposing a cooperative lane-changing strategy with exclusive lanes for CAVs. Du et al. (2018) demonstrated mobility improvement at an unsignalized intersection while proposing a hierarchical coordination strategy for CAVs to traverse through multiple signal-free intersections. Recently, based on V2X communication, Ma et al. (2021) proposed a cooperative adaptive cruise control for CAVs to improve throughput at signalized intersections. However, all the aforementioned studies focus on the safety or mobility implication of CAVs at a local level (i.e., intersections and road segments), without considering the potential network effects of traffic rerouting. Research on the implication of safety and mobility with CAVs at a network level is still limited. Hasibur Rahman and Abdel-Aty (2021) are among the first to investigate the safety and mobility impact of CAVs at a network level with both V2V and I2V technologies to communicate with surrounding vehicles and traffic signals. However, this study relies on microscopic simulation, which may not be easily generalized.

### 2.2.2 Adaptive Decision Making with Information Updates

Studies have investigated the transportation system dynamics in the framework of simulation-based Dynamic Traffic Assignment (DTA) (Antoniou et al., 2011; Gao, 2012; Gao and Chabini, 2006;

Gao and Huang, 2012; Hasibur Rahman and Abdel-Aty, 2021; Lin et al., 2008; Ma et al., 2016; Mahmassani, 2001; Sundaram et al., 2011). Some studies have explicitly considered the impacts of real-time information updates using iterative approaches to account for drivers' experience and learning processes, which may not lead to convergence. For example, Gao (2012) proposed a fixed-point formulation for the user equilibrium with real-time information, which was solved by a heuristic method of successive averages. Based on the space-time network, Ma et al. (2016) proposed an agent-based optimization modeling framework to determine information provision, which can be solved by integrating Lagrangian relaxation-based heuristics and a mesoscopic DTA simulator. Although simulation-based DTA approaches can characterize system dynamics in detail, it may be challenging to calibrate the simulation parameters for a large network. In addition, rigorous mathematical properties (e.g., existence, uniqueness, and convergence) and structural analyses of system interaction may be challenging to achieve (Rambha et al., 2018).

To mitigate the drawbacks of simulation-based DTA approaches, another stream of literature adapts the classic traffic equilibrium notions (e.g., Wardrop user equilibrium (Beckmann et al., 1955; Wardrop, 1952a), stochastic user equilibrium (SUE) (Daganzo and Sheffi, 1977), and DTA (Friesz and Han, 2019)) to capture the impacts of information updates. For example, Acemoglu et al. (2018) generalized classic Wardrop user equilibrium and proposed an information-constrained Wardrop equilibrium model to study the impact of information on traffic congestion. Studies also investigated the route choice behavior and traffic equilibrium patterns in the context of variable message signs (VMSs) (Ban et al., 2009; Lam and Chan, 1996; Li et al., 2016) and advanced travelers information system (ATIS) (Hall, 1996; Henn and Ottomanelli, 2006; Levinson, 2003; Yang, 1998; Yang et al., 1993) using network modeling approaches. For instance, Li et al. (2016) proposed a stochastic network equilibrium model to determine the optimal location of VMS to share travel time information. Yang (1998) proposed a mixed equilibrium assignment model, where informed drivers and uninformed drivers choose routes in user-optimal and stochastic-equilibrium



manners, respectively. However, these studies do not consider the adaptive routing of vehicles.

Du et al. (2014) and Du et al. (2015) proposed equilibrium routing decision (ERD), an extension of SUE at an individual level and in a shorter time frame. However, ERD treats decision making at each time step as independent of each other and does not consider the network impacts of rerouting. Unnikrishnan and Waller (2009), Boyles and Waller (2010), Boyles and Waller (2011), and Rambha et al. (2018) studied user equilibrium with recourse (UER) in an uncertain transportation network where travelers adjust their travel routes depending on en-route information updates. Calderone and Sastry (2017) leveraged the potential function to represent the Markov decision process routing games in the context of ride-sourcing services, where drivers consider the routing over an entire time horizon with real-time information updates. Most of the above-mentioned research on adaptive decision making depicts information implicitly in routing policies, which is challenging for further information systems design and analyses. In addition, the number of routing policies grows exponentially with both uncertain scenarios and paths, which brings significant challenges for real-world applications. To mitigate these issues, we propose a novel and scalable equilibrium-based network model based on the concept of non-anticipativity in stochastic programming (e.g., (Rockafellar and Wets, 1976)) to characterize the equilibrium traffic patterns considering adaptive decision making with information updates.

### **2.3 Methodology**

The proposed research aims to investigate the safety and mobility implications of CAVs and I2V technologies considering en-route information updates and rerouting. Note that en-route rerouting decision does not necessarily imply a dynamic traffic state. En-route rerouting decision could also be because a (static) transportation state is not known to travelers until they pass by information nodes. In this paper, we focus on static traffic patterns.

### 2.3.1 Network Modeling

We use network modeling techniques to capture the interdependence between traffic flow at different links and formulate the adaptive decision making of CAVs when they pass by the I2V infrastructure and receive information updates. Depending on the locations of I2V infrastructure and the travel paths of CAVs, different CAVs may receive different information. Similar to the classic traffic assignment (Wardrop, 1952a), the link travel time and the route choices of all CAVs are coupled. To describe the traffic flow pattern given information updates, we propose a novel computationally tractable static-equilibrium-based formulation of traffic assignment with adaptive routing decision making. In this subsection, we will start with a brief summary of the classic Beckmann's user equilibrium (UE) models (Beckmann et al., 1955). Then, we discuss how to generalize the classic user equilibrium models to reflect adaptive decision making with information updates.

#### 2.3.1.1 Beckmann's Models for UE

We denote a transportation network as a directed graph  $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ , where  $\mathcal{N}$  is the node set (indexed by  $n$ ) and  $\mathcal{A}$  is the set of links (indexed by  $a$ ). Furthermore, we define the OD set  $\mathcal{C} \doteq \mathcal{R} \times \mathcal{S}$  (indexed by  $rs$ ), where  $\mathcal{R}, \mathcal{S} \subseteq \mathcal{N}$  are sets of origin and destination nodes. The route (path) set for connecting an OD pair  $rs \in \mathcal{C}$  is denoted as  $\mathcal{P}_{rs}$ . By Wardrop first principle (Wardrop, 1952a),  $\forall p \in \mathcal{P}_{rs}$  with positive path flow (i.e.,  $x_p > 0$ ),

$$c_p(x) = \min_{q \in \mathcal{P}_{rs}} c_q(x) \quad (2.1)$$

where  $x$  is the path flow vector and  $c_p(x) = \sum_{a \in p} t_a v_a$ , where  $c_p(\cdot)$  and  $t_a(\cdot)$  is the path and link travel costs functions, respectively.

Based on the definition of Wardrop user equilibrium, Beckmann et al. (1955) proposed an equiv-

alent convex mathematical formalization as (2.2) to facilitate the calculation of link traffic flow  $v$  for large-scale transportation networks.

$$\underset{v}{\text{minimize}} \quad \sum_{a \in \mathcal{A}} \int_0^{v_a} t_a(z) dz \quad (2.2a)$$

$$\text{subject to} \quad v \in \mathcal{X}_v, \quad (2.2b)$$

where  $\mathcal{X}_v$  is the projection into the arc flows space of feasible flow set  $\mathcal{X} =: \{(v, x)\}$ .

However, the definition of Wardrop equilibrium and the corresponding formulation and properties are limited to the deterministic case. Stochastic user equilibrium, pioneered by Daganzo and Sheffi (1977), was proposed to model one-stage routing decision making considering perceived uncertainties of travel time. Travelers navigate through the network based on their perceived travel time and there are no en-route information updates. Therefore, the routing decision making of travelers is made at the beginning of the trip, i.e., single-stage decision making. However, in our proposed model, travelers know the probability distribution of travel times a priori and could receive information updates on the actual realization if they pass information nodes.

### 2.3.1.2 Adaptive Traffic Assignment Models

To distinguish from the classic (one-stage) stochastic user equilibrium notion (Daganzo and Sheffi, 1977), we denote our traffic equilibrium as stochastic user equilibrium with recourse (SUER) in this study, as formally defined in Definition 1. We assume that drivers are rational and risk-neutral, and have a common prior belief about the link travel time. The prior belief is measured by the probability distribution of system uncertain parameters  $\xi$ , which can be estimated from historical traffic data. I2V devices will provide information updates on  $\xi$  when CAVs pass through I2V nodes, denoted as  $\mathcal{N}_I$ . We further assume that information nodes will provide true and consistent

information on system uncertain parameters  $\xi$ , so that for the first time a driver passing through node  $n \in \mathcal{N}_I$ , he/she will receive a specific realization of  $\xi$ , which will be taken into account for their future routing decisions. This setting indicates that there are at most two possible stages in which a traveler could be in, namely before and after receiving information. In the first stage (i.e., before they receive information updates), travelers make routing decisions based on the expectation of the path travel time given prior knowledge of the probability distributions of link performance functions. In the second stage (i.e., after they receive information updates), travelers adjust their routes based on the new information. Travelers will receive information when they pass an information node and we do not assume travelers need to receive all information at the same time. Note that our traffic equilibrium aligns with the user equilibrium with recourse (UER) notion originally proposed in (Unnikrishnan and Waller, 2009). However, instead of using a policy-based modeling approach, we adopt a completely different modeling strategy inspired by two-stage stochastic programming with recourse. We note that since the proposed modeling framework is based on two-stage stochastic programming concepts, the proposed modeling strategies can be naturally extended to a multi-stage SUER case, where information nodes reveal partial information about the systems. To keep the focus clear, we left the treatment of multi-stage SUER for the future.

**Definition 1.** (*two-stage SUER*). *In two-stage SUER, the following two conditions hold:*

- *the expected travel times/costs on all paths used in the first stage are equal, and less than those which would be experienced by a single vehicle on any unused path;*
- *the travel times/costs on all paths used in the second stage are equal, and less than those which would be experienced by a single vehicle on any unused path.*

Notice that the drivers may adjust their travel routes after receiving information updates. Therefore, “paths selected in the first stage” should bundle all the possible paths that a driver can still choose

after they receive information. More formally, we define the “bundled paths” as *hyperpath*, as shown in Definition 2.

**Definition 2.** (*hyperpath*). A hyperpath, denoted as  $\mathcal{P}_k^{rs}$ , is the  $k^{\text{th}}$  set of paths that connect the same OD pair  $rs$  and share the same sequence of links from origin  $r$  to the first information node  $i$  ( $i = s$  if there are no information nodes on a path).

Inspired by the mathematical formulation of Wardrop equilibrium proposed by Beckmann et al. (1955), we construct a convex optimization problem, whose optimal solutions are consistent with our definition of two-stage SUER. The convex optimization model that can generate the equilibrium outcome according to Definition 1 is presented in model (2.3).

$$\min_{x_p(\xi) \geq 0, \forall p, \xi} \mathbb{E}_\xi \sum_{a \in \mathcal{A}} \int_0^{v_a(\xi)} t_a(u, \xi) du \quad (2.3a)$$

$$\text{s.t.} \quad v_a(\xi) = \sum_{rs \in RS} \sum_{p \in \mathcal{P}^{rs}} \delta_{ap} x_p(\xi), \forall a \in \mathcal{A}, \xi \in \Xi \quad (2.3b)$$

$$(\gamma^{rs}(\xi)) \quad \sum_{p \in \mathcal{P}^{rs}} x_p(\xi) = q^{rs}, \forall rs \in RS, \xi \in \Xi \quad (2.3c)$$

$$(\lambda_{a,k}^{rs}(\xi)) \quad \sum_{p \in \mathcal{P}_k^{rs}} \delta_{ap}^+ x_p(\xi) = x_a^{rs}, \forall rs \in RS, a \in \mathcal{A}, \xi \in \Xi \quad (2.3d)$$

where:

$v_a$  : traffic flow on link  $a$ ;

$t_a(\cdot, \cdot)$  : travel time function of link  $a$ ;

$q^{rs}$  : travel demand from  $r$  to  $s$  (model input);

$x_p^{rs}(\xi)$  : traffic flow on path  $p$  that connects  $r, s$  at scenario  $\xi$ ;

$x_a^{rs}$  : traffic flow on link  $a$  from  $r$  to  $s$  that haven't received information on  $\xi$ ;

$\delta_{ap}$  : link-path incidence scaler, which equals to 1 if link  $a$  belongs to path  $p$ , and 0 otherwise;

$\delta_{ap}^+$  : link-path incidence scaler, which equals to 1 if link  $a$  belongs to path  $p$  and has not passed through any node  $n \in \mathcal{N}_I$ , and 0 otherwise;

$\mathcal{P}_k^{rs}$  : The path set for those paths share the same sequence of links before they receive uncertainties information, with  $k$  denote the  $k^{\text{th}}$  set for  $rs$ ;

$\gamma, \lambda$  : dual variables of corresponding constraints.

The basic idea of the model (2.3) is that we first relax path flow  $x_p$  and links flow  $v_a$  to be scenario dependent, and then we enforce non-anticipativity constraint (2.3d) to guarantee that the paths flow is measurable according to the uncertainty set, i.e. the flow on the path segments before receiving information (i.e.,  $\delta_{ap}^+ = 1$ ) should not be measurable by  $\xi$ . Constraint (2.3b) aggregates the path flow to each link for each scenario  $\xi$ ; constraint (2.3c) restricts the summation of path flow connecting each OD pair should equal to OD demand. The objective function (2.3a) is constructed in this way so that the optimal solutions of the optimization problem (2.3) satisfies the two conditions of two-stage SUER, as stated in Definition 1. This result is stated more formally in Theorem 2.3.1.

**Theorem 2.3.1.** *(two-stage SUER) The traffic flow pattern is following two-stage SUER principles (i.e., Definition 1) if and only if it is the optimal solution to the optimization problem (2.3).*

*Proof.* See A.1. □

The proposed formulation is intuitive in the sense that the final formulation, i.e., model (2.3), closely relates to the classic UE formulation. However, the rigorous proof of the equivalence between the solutions of the proposed model (2.3) and the traffic patterns following the definition of two-stage SUER (i.e., Definition 1) is non-trivial and novel. As far as the authors are aware, no other work has utilized a stochastic programming concept to model a UER problem, which could lead to significant computational advantages since the proposed models belong to multi-

stage stochastic programming problems, for which a large body of literature has been dedicated to developing efficient computational algorithms.

In addition, we note that there are two other well-known modeling frameworks, DTA and Markovian traffic equilibrium (MTE), that could be relevant for the studied problem, but are not adopted in this study. The reasons are explained as follows.

On the one hand, we adopt static-equilibrium-based models instead of DTA for three main reasons. First, we focus on the stochastic nature of the transportation system instead of its dynamic nature. In our setting, we assume link performance functions are uncertain, whose realizations can only be known when travelers pass by certain information nodes. In other words, the realizations of link performance functions are static (do not depend on time), although the link performance functions themselves are uncertain (unknown before reaching information nodes). Second, while DTA could be used to model en-route rerouting with uncertainties, it is computationally costly, while our proposed method is a convex optimization, as shown in Section 2.3.1.2. Third, static approaches have been demonstrated to be effective to represent the adaptive routing decision of travelers in transportation networks (Unnikrishnan and Waller, 2009).

On the other hand, although the fundamental concept of MTE appears relevant, the current mathematical framework of MTE (e.g., (Baillon and Cominetti, 2008)) may not be suitable for studying the proposed problem. This is because the MTE formulation focuses on the variation of travelers' perception of link travel time (Baillon and Cominetti, 2008). In other words, the stochasticity in MTE refers to the stochasticity in travelers' perception of the link travel time, rather than to the stochasticity of the link performance function itself. This is an important distinction because following the MTE setting, one can adopt discrete choice models to model the traffic distribution at any given node. However, in the case when the link performance function is uncertain, travelers make decisions considering expected travel time based on prior knowledge and information

updates on the probability distributions and the discrete choice model is not applicable.

### 2.3.1.3 Illustrative Example for Two-stage SUER

To illustrate two-stage SUER, we use a simple four-node network, as shown in Figure 2.1.

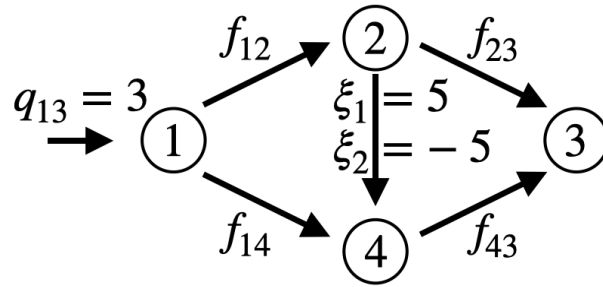


Figure 2.1: Four-node Network with Stochastic Link Cost

In Figure 2.1, there are 3 units of travel demand from node 1 to node 3. The link cost, as a function of link flow  $f$ , is shown on each link. The travel cost of link 2-4, denoted as a random variable  $\xi$ , is uncertain, which could be influenced by random events, such as traffic accidents/incidents, weather, tolls, recommended speed, etc. We assume  $\xi$  has two realizations, 5 and -5<sup>1</sup>, with equal probability in this toy example. Node 2 is the only information node that reveals the realization on  $\xi$  to the drivers passing by node 2. From node 1 to node 3, there are three paths:  $p_1 = \{1, 2, 3\}$ ,  $p_2 = \{1, 2, 4, 3\}$ ,  $p_3 = \{1, 4, 3\}$ . There are two possible hyperpaths in the first stage:  $\mathcal{P}_1^{rs} = \{p_1, p_2\}$  and  $\mathcal{P}_2^{rs} = \{p_3\}$ . In the first stage, drivers will decide whether to take link 1-2 or 1-4, considering the historic probability distribution of  $\xi$ . If drivers choose link 1-2 in the first stage, they will have an option later to decide whether to go on route 2-4-3 or route 2-3 based on the new information

<sup>1</sup>We select 5 and -5 to simplify the verification of solutions without losing generality. Negative travel costs are possible given the possibility that drivers may have incentives (e.g., scenery) to use certain routes.



received at node 2; while drivers who choose link 1-4 will not receive information updates during their travel and will following path 1-4-3.

The optimal solutions of model (2.3) are shown in Table 2.1. We can see that the traffic flow on the hyperpaths  $\mathcal{P}_1^{rs}$  and  $\mathcal{P}_2^{rs}$  are independent of scenario parameter  $\xi$ . This is because of non-anticipativity constraints so that drivers are not able to make scenario-dependent routing decisions before vehicles receive information. But after vehicles receive information updates, the traffic will redistribute between different paths in each hyperpath depending on scenario  $\xi$ . For example, we can see that all traffic (7/3 units) in hyperpath  $\mathcal{P}_1^{rs}$  select path  $p_1$  in scenario  $\xi_1$  and they select path  $p_2$  in scenario  $\xi_2$ . From the (expected) costs, as presented in Table 2.1, we can verify that: (1) Hyperpath  $\mathcal{P}_1^{rs} = \{p_1, p_2\}$  and  $\mathcal{P}_2^{rs} = \{p_3\}$  have the same expected cost <sup>2</sup>; and (2)  $p_1$  has lower second-stage travel cost than  $p_2$  when  $\xi = 5$  so that  $p_1$  is selected while  $p_2$  is unused in the second stage for this scenario. Opposite observations hold when  $\xi = -5$ . Therefore, the solutions of model (2.3) indeed satisfy Definition 1.

The traffic patterns reported in Table 2.1 are dramatically different from the classic user equilibrium solutions with expected link costs on link 2 – 4, where  $x_{p_1} = x_{p_3} = 1.5$ , and  $x_{p_2} = 0$ . The interpretation for this difference is that because I2V is available at node 2, CAVs expect to receive information updates if they travel through  $\{p_1, p_2\}$  in the first stage. Therefore, more CAVs are attracted to choose  $\{p_1, p_2\}$  compared to  $\{p_3\}$  in a two-stage SUER in contrast to the user equilibrium solutions under Wardrop first principle.

We also note that equilibrium traffic patterns may also be sensitive to where information is shared. For example, if information on  $\xi$  is shared at node 1, all drivers will have perfect information about link cost of 2 – 4 at the beginning of the trips. Therefore, the traffic equilibrium solutions

---

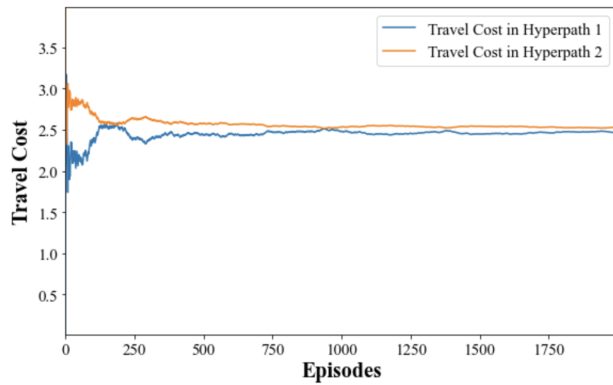
<sup>2</sup>Note that although the expectation of travel costs of both  $\mathcal{P}_1^{rs}$  and  $\mathcal{P}_2^{rs}$  are the same,  $\mathcal{P}_1^{rs}$  has a higher variance of travel costs compared to  $\mathcal{P}_2^{rs}$ . Therefore, for those travelers value reliability of travel costs, more traffic will be shifted from  $\mathcal{P}_1^{rs}$  to  $\mathcal{P}_2^{rs}$ . This is beyond the scope of this research and will be left for future investigation.

are scenario dependent with  $x_{p_1} = x_{p_3} = 1.5$ ,  $x_{p_2} = 0$  at scenario  $\xi_1$  and  $x_{p_1} = x_{p_3} = 0$ ,  $x_{p_2} = 3$  at scenario  $\xi_2$ .

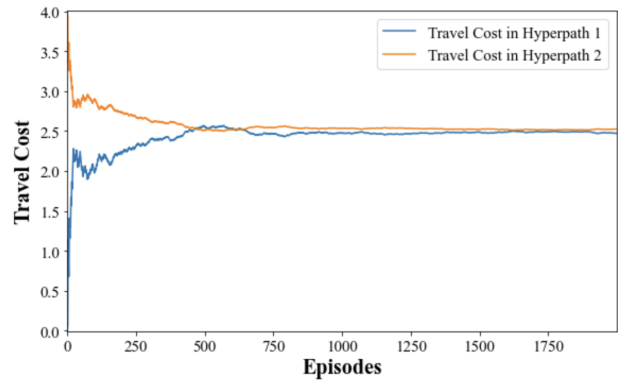
Table 2.1: Equilibrium solutions with risk-neutral adaptive behaviors

Path	Flow		Travel Costs			
	$\xi_1$	$\xi_2$	$\xi_1$	$\xi_2$	Exp.	Var.
$p_1$	7/3	0	14/3	-	2.5	9.4
$p_2$	0	7/3	-	1/3	2.5	9.4
$p_3$	2/3	2/3	4/3	11/3	2.5	2.7

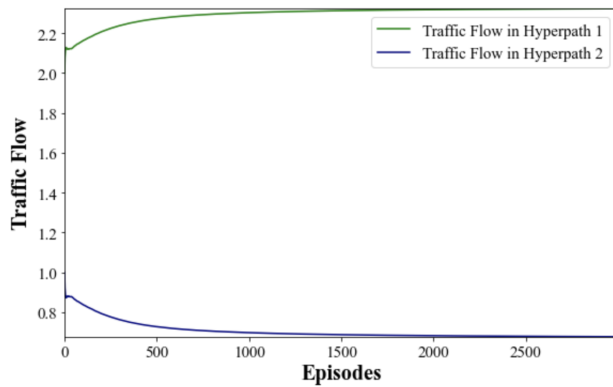
To demonstrate the traffic patterns will be indeed stabilized given that travelers will update their prior knowledge on the probability distribution of link performance functions and receive en-route information updates, we have simulated the travelers' en-route rerouting behaviors in response to uncertainties and information updates using the same example shown in Figure 2.1. We consider two cases. The first case (denoted as the global information case) assumes that travelers will learn the probability distributions for all links regardless of whether they travel on the links or not. This case is plausible considering travelers can get access to the realization of link performance through news or social media afterward. In the second case (denoted as the local information case), we assume that travelers will only learn the probability distributions on the links they travel on. In the simulation, travelers will update their knowledge of probability distributions and select the routes to minimize the expected travel time based on their up-to-date information set. Figures 2.2a and 2.2b demonstrate that the expected travel costs of both hyperpaths will converge to an equilibrium travel costs identical to our model solution in Table 2.1 for both cases. Similarly, the traffic patterns converge to those in Table 2.1, as shown in Figures 2.2c and 2.2d.



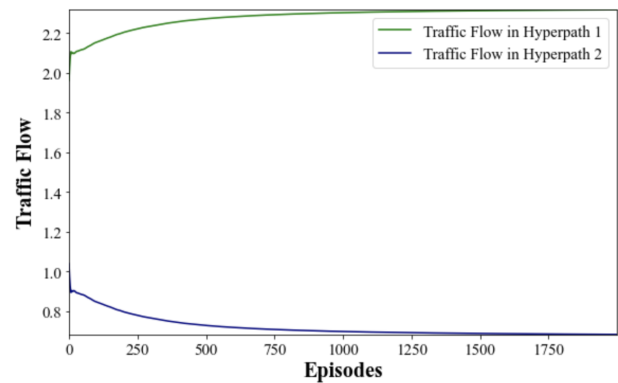
(a) Travel Costs with Global Information



(b) Travel Costs with Local Information



(c) Traffic Flow with Global Information



(d) Traffic Flow with Information

Figure 2.2: Convergence of Traffic Patterns and Expected Travel Time

### 2.3.2 Parametric Estimation

Network modeling and system safety assessment require link-level information (i.e., traffic link performance function and link crash risk). In this subsection, we will discuss how to estimate the key parameters using measurable data.

### 2.3.2.1 Link Performance Function

Researchers have used different link performance functions in the past (Branston, 1976). By far, the Bureau of Public Roads (BPR) function is one of the most widely adopted functions because of its simplicity and convenience to represent congestion behavior and free-flow travel time. We adopt the BPR function in this paper. The BPR functions are expressed as Equation (2.4).

$$t = t_0 \left[ 1 + \alpha \left( \frac{v}{c} \right)^\beta \right] \quad (2.4)$$

where  $t_0$ ,  $\alpha$ ,  $\beta$ , and  $c$  are parameters that define the shape of the BPR function and needed to be estimated for different types of roadways. The travel time will increase with the ratio of the flow  $v$  and the capacity  $c$  to represent congestion effects.

In a stochastic environment, parameters  $t_0$ ,  $\alpha$ ,  $c$ , and  $\beta$  in link performance function (2.4) could be random variables. In this paper, we assume that with an increasing amount of traffic data, parameters in link performance functions can be categorized as probability distributions. The uncertainties could be due to recurring events that will affect the link travel time, such as weather events, construction/maintenance activities, traffic accidents, etc. Travelers will learn the probability distribution of link performance functions through day-to-day learning. While the link performance functions are uncertain, the realization of uncertainties in a given scenario does not change during the study period, thus warranting the stability of the traffic. The link performance function at a specific scenario  $\xi \in \Xi$  can be expressed in Equation (2.5).

$$t_\xi = t_{0,\xi} \left[ 1 + \alpha_\xi \left( \frac{v_\xi}{c_\xi} \right)^{\beta_\xi} \right] \quad (2.5)$$

where  $\xi = \{t_{0,\xi}, \alpha_\xi, c_\xi, \beta_\xi\}$  are random parameters. At the beginning of the trips, drivers have a

prior belief about the probability distribution of the random parameters, which could be updated along the CAVs' travel paths if I2V information sharing is available. Note that, without loss of generality, the information that CAVs receive from I2V infrastructure can be an updated probability distribution of these random parameters instead of a specific realization.

Link travel time  $t$  will depend on the uncertain scenario  $\xi$  in addition to link flow  $v$ . To estimate the uncertain parameters for a specific scenario, one can use non-linear least squares regression. The input data for the regression model is  $(t, v)$ , which could be simulated in a microsimulation platform when they are not able to be directly measured from real-world traffic sensors.

### 2.3.2.2 Collision Risk Function

Collision risks can be defined in multiple ways. When crash data is available, collision risk can be measured as the probability of collision happening under specific traffic conditions. In this section, we will present the methodology to calculate collision risk. The calculated collision risk will be further used to estimate the relationship between collision risk and link traffic flow.

When real-world collision data is available, we encode the collision risk using a binary variable  $Y$ , with 1 indicating collision happens and 0 indicating no collision. Since  $Y$  is a binary variable, we propose a discrete choice-based crash risk model to formulate the probability of a collision  $p$ . The discrete choice-based crash risk model is also adopted by Shi and Abdel-Aty (2015) and Hasibur Rahman and Abdel-Aty (2021).

The target variable is a binary indicator  $Y$  indicating whether collision happens or not with probability  $p$ . In other words,  $Y$  is a Bernoulli trial and can be expressed as a probability distribution (2.6).

$$Y \sim \text{Bernoulli}(p) \quad (2.6)$$

Discrete choice model is able to estimate the probability  $p$ , indicating that with a probability of  $p$  collision may happen (i.e.,  $Y = 1$ ) and with a probability of  $1 - p$  otherwise. The formulation of the logit model is shown in Equation (2.7) for any given link.

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right) = \beta_0 + X\beta_1 \quad (2.7)$$

where:

$p$  : the likelihood of collision happens

$\beta_0$  : the constant

$X$  : the explanatory variables, such as traffic flow  $q$

$\beta_1$  : the coefficients of the explanatory variables

While explanatory variables  $X$  can include different variables, in this study, we focus on traffic flow  $q$  and treat other factors as exogenous variables, whose impacts will be included in link-specific constant coefficient  $\beta_0$ . Therefore, the collision risk for each link is shown in Equation (2.8).

$$CR^{\text{Link}} = p = \frac{e^{\beta_0 + \beta_1 v}}{1 + e^{\beta_0 + \beta_1 v}} \quad (2.8)$$

**Remark 1:** Since traffic flow is one of the most important factors influencing traffic safety (?), we characterize the collision risk of a link (equation (2.8)) as a function of the traffic flow of that link. However, we note that our proposed model can be extended to incorporate the potential correlation of link safety risk functions among neighboring links for different scenarios. For example, in a

given scenario  $\xi$ , equation (2.8) can be updated as  $CR^{\text{Link},\xi} = p = \frac{e^{\beta_{0,\xi} + \beta_{1,\xi}q}}{1 + e^{\beta_{0,\xi} + \beta_{1,\xi}q}}$ . This can offer the flexibility to model the link collision risk considering the effect of neighbor links because the parameters for neighbor links can be specified for each scenario to consider their correlation. Since the focus of this paper is to propose a new network model to study the traffic equilibrium with en-route information updates, we will leave the estimation and implementation of the neighboring effects for the future.

**Remark 2:** Scenario probability and crash risk are interdependent. But there are three main reasons that we did not consider the endogeneity of scenario probability in this modeling framework. First, the scenario we consider is not solely based on whether accidents occurred or not. The link performance functions in each scenario are determined by multiple factors, including weather events, traffic management, and/or accidents. The probability distribution of the link performance function is learned from the travel experience of travelers. Therefore, the probability of collision does not equal the probability of a certain scenario. Since accidents are only part of the factors that define the scenarios, in this manuscript, we treat the scenario probability as exogenous. Second, accidents are rare events and the probability of accidents is usually low and difficult to measure directly using a probability distribution function. In such cases, only focusing on accidents to represent each scenario may not be practical in a real-world setting. Third, if scenario only represents accidents and we can measure accident probability, we can use iterative approaches to find the fixed point of the following function  $x = f(x)$ , where  $x$  is the probability vector of scenarios (i.e.,  $pr(\xi)$ ) and  $f(\cdot)$  is the mapping from  $pr(\xi)$  to crash risk, which combines model (2.3) and equation (2.8). However, we note that the convergence of such iterative approach may not be guaranteed for high-dimensional scenarios.

### 2.3.3 Network Safety Analyses

In order to conduct an analysis on the impact of information sharing locations on network safety, we need to mathematically formulate different information sharing strategies and embed them in the two-stage SUER model (2.3), and quantify the network safety metrics. This subsection focuses on these two aspects.

#### 2.3.3.1 Representation of Information Sharing Locations

Denote the set of I2V information sharing locations as  $\mathcal{I} \subset \mathcal{N}$ . When CAVs passed by these locations, CAVs will receive information updates for their future routing decision making.

Denote information node indicator variable  $y_n, \forall n \in \mathcal{N}$ .  $y_n = 1$  if I2V infrastructure is placed at node  $n$ ; and  $y_n = 0$  otherwise. Denote a path that consists of  $K$  links, as  $p = \{a_1, a_2, \dots, a_K\}$ .  $\delta_{ap}^+$  in Model (2.3) can be calculated using Equation (2.9).

$$\delta_{ap}^+ = \prod_{n \in \mathcal{N}_{ap}} (1 - y_n) \quad (2.9)$$

where  $\mathcal{N}_{ap}$  is the node set on path  $p$  before reaching to link  $a$ .

Using Figure 2.1 as an example, if I2V infrastructure is placed at node 2,

$$y_n = \begin{cases} 1, & \text{if } n = 2 \\ 0, & \text{otherwise.} \end{cases} \quad (2.10)$$

For path  $p_2 = \{1, 2, 4, 3\}$  and link  $a = \{(4, 3)\}$ , we have  $\mathcal{N}_{ap} = \{1, 2, 4\}$ . Therefore,  $\delta_{ap}^+ = \prod_{n \in \mathcal{N}_{ap}} (1 - y_n) = 0$ . In other words, CAVs traveling on path  $p_2$  have received the information



before they travel on link  $\{(4,3)\}$ .

### 2.3.3.2 Network Safety Assessment

Given crash risk  $CR$  at link level as a function of link traffic flow  $q$  (i.e., Equation (2.8)), we can derive the collision risk (CR) for the network as the total risk experienced by all the vehicles in the network. In a stochastic environment, the parameters  $\beta_0/\beta_1$  in Equation (2.8) for any link  $a$  will be random variables. We denoted the collision risk for link  $a$  at scenario  $\xi$  as  $CR_{a,\xi}^{\text{Link}}$ . Collision risk for the network at a specific scenario  $\xi \in \Xi$  can be expressed in Equation (2.11).

$$CR_{\xi}^{\text{Network}} = \sum_{a \in A} CR_{a,\xi}^{\text{Link}} v_{a,\xi} \quad \forall \xi \in \Xi \quad (2.11)$$

where  $v_{a,\xi}$  is the link flow at link  $a$  in scenario  $\xi$ .

In this study, the network safety indicator is measured as the expected network safety risks. We seek to evaluate the impacts of providing information at different locations on the expected network safety risks,  $CR^{\text{Network}}$ , which can be calculated in Equation (2.12).

$$CR^{\text{Network}} = \mathbb{E}_{\xi} \{CR_{\xi}^{\text{Network}}\} \quad (2.12)$$

## 2.4 Numerical Results

We tested our proposed methodologies on the Orlando transportation network, as shown in Figure 2.3, which serves the traffic between three major attractions in Orlando metropolitan area, which are Orlando International Airport (Node 17), Universal Studio (Node 11), and Disney world (Node 3). The network covers an area of 78.87 mile<sup>2</sup>. This network has a total of 27 two-way links,

which comprise both freeways and multi-lane highways. The archived traffic data, collected by the Regional Integrated Transportation Information System (RITIS) from February 1<sup>st</sup>, 2020 to April 30<sup>th</sup>, 2020, contains speed, traffic volume, and traffic occupancy. The crash records during the studied time period were collected from Florida Highway Safety and Motor Vehicles. Based on these raw data, we are able to estimate the link flow, link travel time, and link crash risks at each hour, which will be used to estimate the link performance functions and link crash risk functions, as discussed in Section 2.3.2. In this section, we demonstrate the impacts of different information sharing strategies on network traffic distribution, mobility, and safety. For the base case, incident links include 14-17, 17-14, 3-5, and 5-3, which are four links with relatively higher collision risks. Each incident link could be in two scenarios: *normal scenario* and *incident scenario*. We differentiate the *incident scenario* from *normal scenario* by reducing the incident links' capacity by 50% in Equation (2.5). The capacity reduction could be due to incidents such as traffic accidents, weather events, and road condition deterioration. We consider information sharing at node 10 (a central node in the network) as the base case if not further specified.

In the remainder of this section, we first analyze the network mobility and safety for the base case; then we conduct sensitivity analyses on key parameter settings (including information sharing locations, incident severity, and OD demand).

#### 2.4.1 Base Case Results

Both traffic congestion level and network collision risk could be different under different information sharing strategies due to rerouting. Travelers' routing decisions can be changed as en-route information is received. To demonstrate this, we show the traffic flow changes for O-D from node 17 (Orlando Airport) to node 3 (Disney World) in Figure 2.4, which depicts the change in link traffic between *normal scenario* and *incident scenario* based on the information they receive at

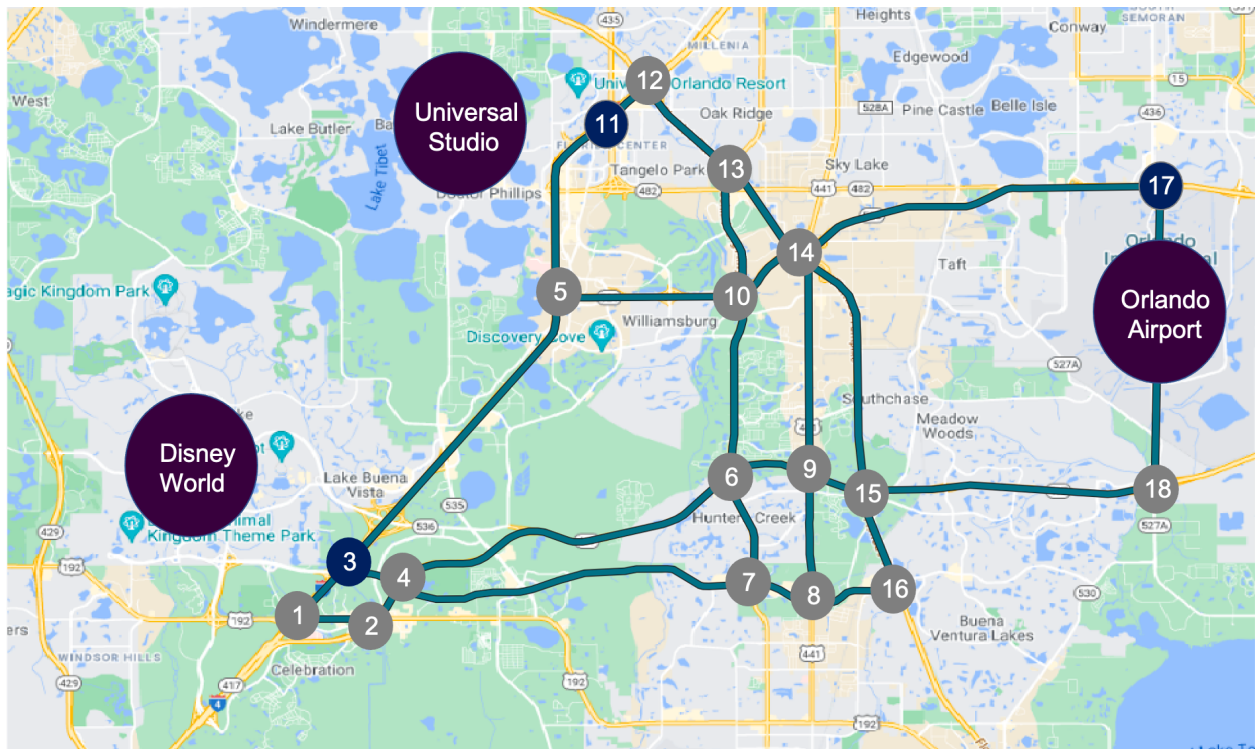


Figure 2.3: Study Area in Orlando

node 10. When travelers receive information at node 10 indicating that no incident happens in link 5-3 (i.e., *normal scenario*), more traffic will use the path 10-5-3. In contrast, when they receive information at node 10 indicating incident happens (i.e, *incident scenario*), travelers originally take the path 10-5-3 in *normal scenario* will re-route to path 10-6-4-3. The traffic on all the other links does not change because they either do not receive information updates along their routes or do not have better routing options despite knowing the incident occurred.

To quantify the impacts of information sharing on network mobility, we measure congestion using the “v/C” ratio (i.e., link flow / link capacity) of a link. Figure 2.5 shows the relative difference in traffic congestion levels between two cases: (1) information is shared at node 10 and (2) no information is shared. Figure 2.5a shows that in the *normal scenario*, when information is shared,

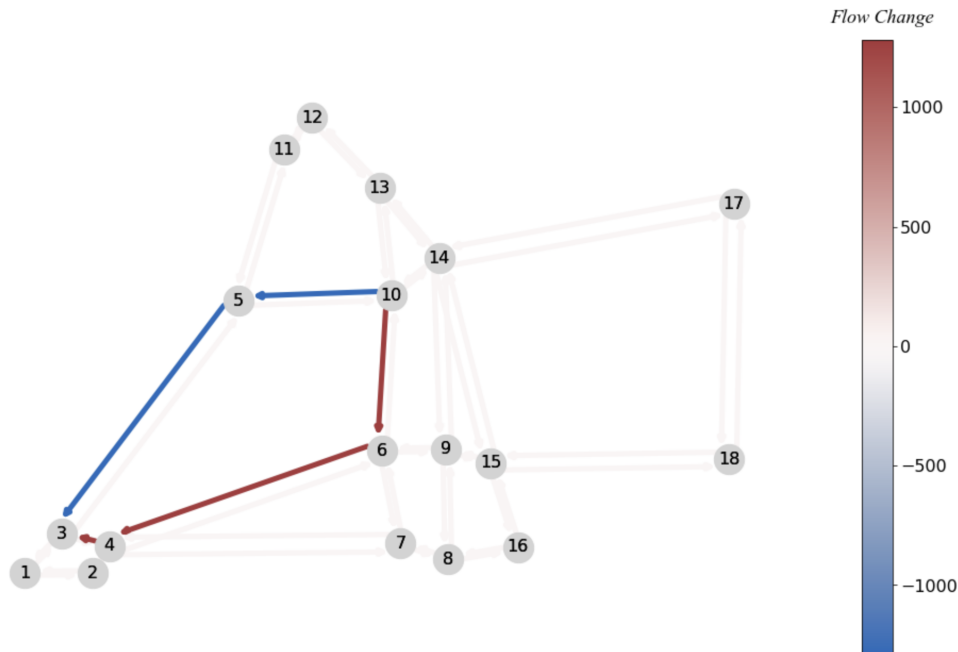


Figure 2.4: Link Flow Changes Between Normal Scenario and Incident Scenario for the Base Case

links 10-5-3 significantly increase the congestion levels, while links 10-6-4-3 reduce the congestion levels, compared with the case when information is not shared. Opposite observation holds for the *incident scenario*. The reason is because for traffic going from node 10 to node 3, their routing will depend on the information they received about the state of link 5-3. On the other hand, for the links before information is received, congestion levels slightly increase in links 17-14-10 while slightly decrease in links 17-18-15-9-6 regardless of the actual realization on the incident links. This is because when node 10 is an information node, more traffic will prefer to pass by node 10 from node 17 so that they can receive information updates to inform their future rerouting decisions. From Figure 2.5, we also observe that information sharing at node 10 mainly affects the traffic routing to node 3. The reason why information sharing does not affect the routing of travelers with destination to other nodes is that those travelers may not have better routing options to avoid incident links despite knowing incidents occurred (e.g., travelers to node 17 prefers link 14-17

regardless of incidents) or they simply do not utilize incident links regardless of which scenario the system is in (e.g., travelers to node 11).

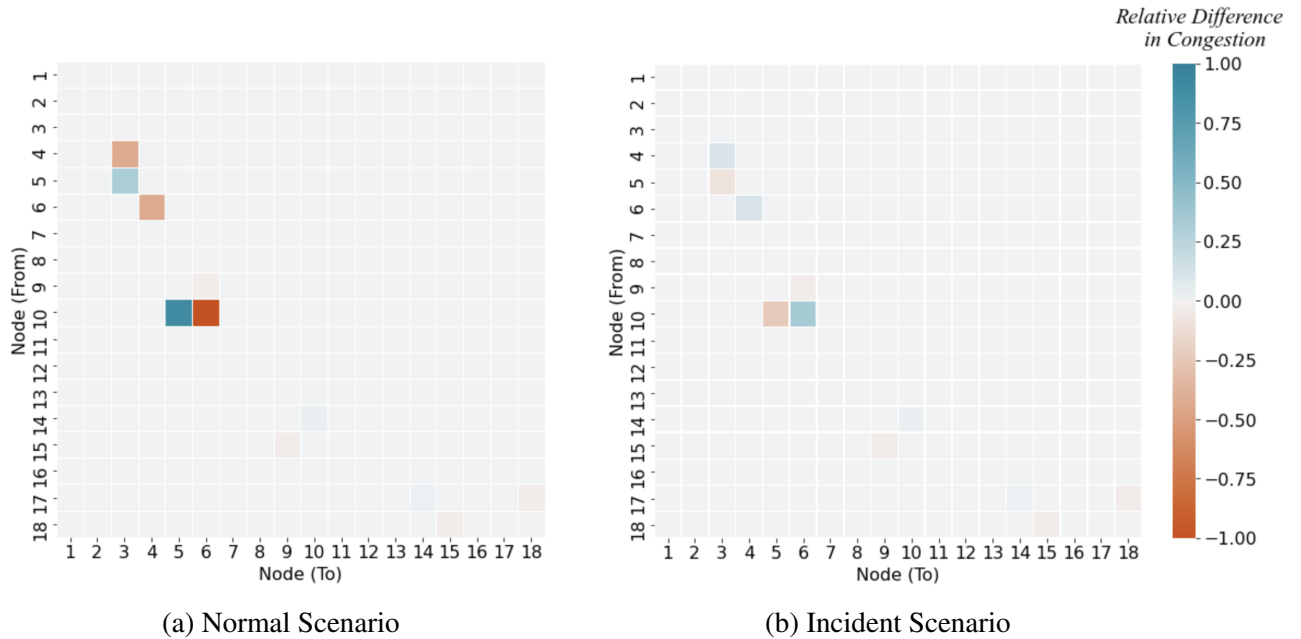


Figure 2.5: Relative Difference of Congestion Level between Information Sharing and No Information Sharing

Traffic rerouting caused by information sharing may further lead to changes in link collision risk over the network. Figure 2.6 illustrates the relative difference of collision risk in each link between two cases: (1) information is shared at node 10 and (2) no information is shared. Figure 2.6a shows that in the *normal scenario*, link 10-5 has higher collision risk, while links 10-6-4-3 have lower collision risk when information is shared compared with when no information is shared. Opposite results for these links are observed in Figure 2.6b for the *incident scenario*. The increase/decrease of collision risk in these links is consistent with the increase/decrease of congestion levels as shown in Figure 2.5b. However, we note that a lower congestion level does not always imply a lower collision risk. For example, link 15-9 has a lower congestion level when information is shared

in both the *normal and incident scenarios*, but this link has a slightly higher collision risk in both scenarios. These results indicate the value of using a network perspective to model traffic redistribution and to gain a better understanding of the impact of information sharing on network traffic safety.

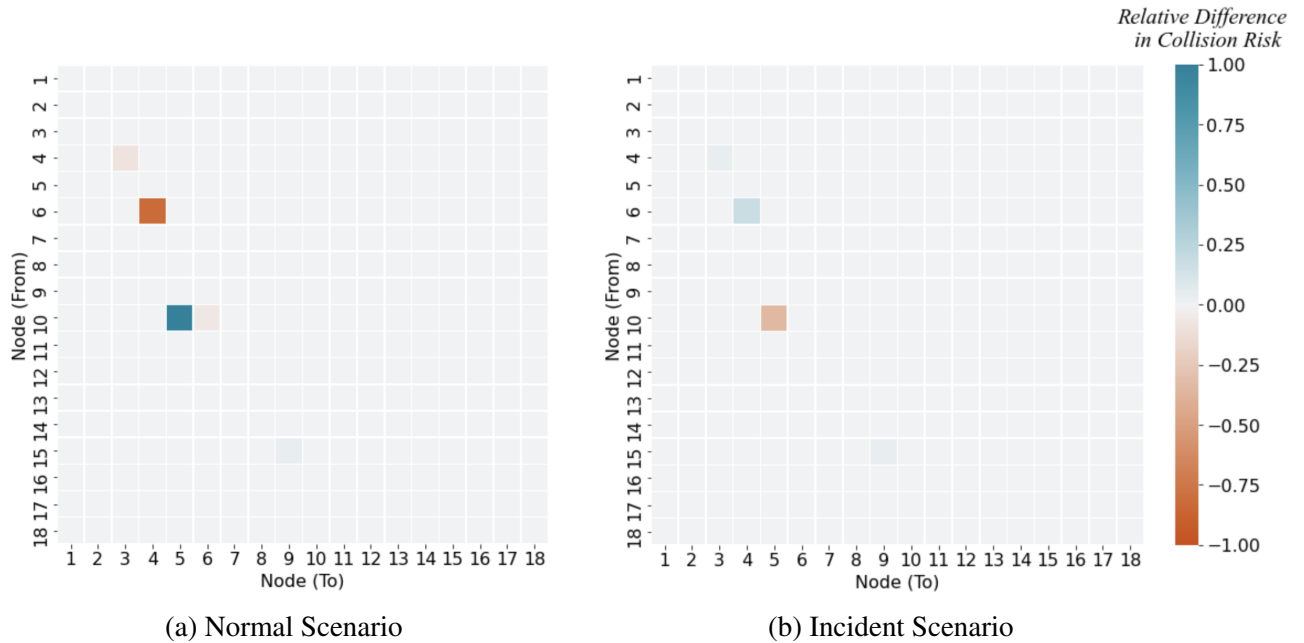


Figure 2.6: Comparison of Collision Risk at each Link between Information Sharing vs. No information Sharing

## 2.4.2 Sensitivity Analyses

### 2.4.2.1 Sensitivity of Information Sharing Strategies

First, we investigate the network effects of information sharing at different individual nodes on network safety and mobility, as shown in Figure 2.7. In Figure 2.7a, we can see that information sharing at nodes 10 or 14 provides the most network benefits in terms of reducing network collision

risk. The reason could be that nodes 10 and 14 are two central nodes in the network and information sharing at these two locations allow traffic passing by (traffic volume 11,240 veh/hour and 13,549 veh/hour at node 10 and 14, respectively) to update their routing decisions, which are beneficial to the whole system. However, one should not conclude that sharing information at nodes with high traffic volume is always better. For example, information sharing at node 4 and node 6, which also have high traffic volume at around 11,255 veh/hour, leads to the highest network collision risks. This again indicates the necessity of using a network modeling strategy to capture the complex interaction among the system.

In Figure 2.7b, we see that information sharing at node 17 leads to the least total travel time in the network. Notice that information sharing at nodes 10 or 14 also benefits the travelers in terms of network mobility. This is because incidents may happen at the links 3-5/5-3 and 14-17/17-14, which are on the shortest path of OD demand 3-17 and 17-3. Information sharing at nodes 10 and 14 allows a large amount of traffic from node 3 (47% of the total incoming traffic at nodes 10/14) and node 17 (50% of the total incoming traffic at nodes 10/14) to choose alternative paths depending on if an incident happens or not. In contrast, information sharing at node 5 is not a good option because, on one hand, traffic from node 3 will be encouraged to utilize link 3-5 to access information, which may experience a delay if an incident happens on that link. On the other hand, traffic from link 10-5 and link 11-5, which is 52% of the total incoming traffic at node 5, do not benefit from receiving information due to a lack of alternative travel routes.

Second, we compare between four information sharing strategies, including (1) *best single node*: information sharing at the best single node (i.e., nodes 10 or 14 for safety and node 17 for mobility), (2) *worst single node*: information sharing at the worst single node (i.e., nodes 4 or 6 for safety and nodes 1,2, 5, 7,8,9,11,12,13,15,16, or 18 for mobility), (3) *perfect information sharing*: information sharing at all the nodes, and (4) *no information sharing*: no information sharing in any nodes, with network safety and mobility results shown in Figure 2.8a and Figure 2.8b, respectively.

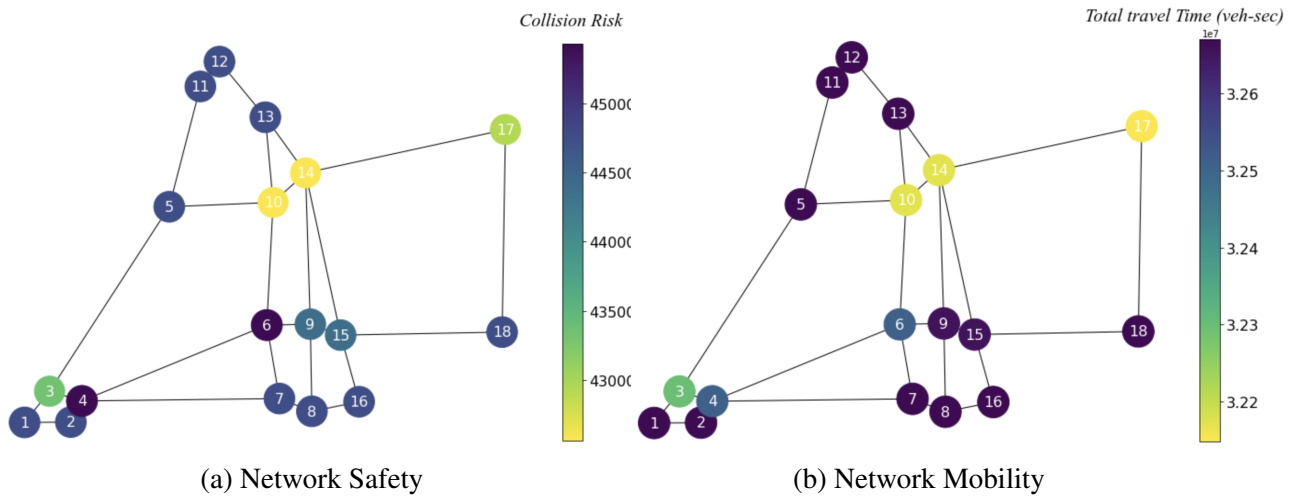


Figure 2.7: Effect of information sharing at individual nodes separately

We can see that *perfect information sharing* is the best strategy for both network safety and mobility. However, we note that no information sharing is not always the worst strategy. For example, in Figure 2.8a, sharing information at node 4 or 6 is the worse strategy compared with no information sharing in terms of network safety. In Figure 2.8b, no information sharing and information sharing at the worst single nodes will lead to the similar total travel time in the network. The reason for this is that information sharing allows for additional travelers rerouting but travelers make rerouting decisions to optimize their own utilities without considering the potential negative externalities for other travelers.

Third, we investigate the Pareto optimal information sharing locations for both network safety and mobility. We assume nodes 3, 6, 10, 11, and 17 are candidate information sharing nodes. Figure 2.9a and 2.9b show the impacts of information sharing for all the subsets of the candidate information sharing nodes for network safety and mobility, respectively. From Figure 2.9a, it can be seen that the network collision risk is the lowest when information is shared at nodes 3 and 10



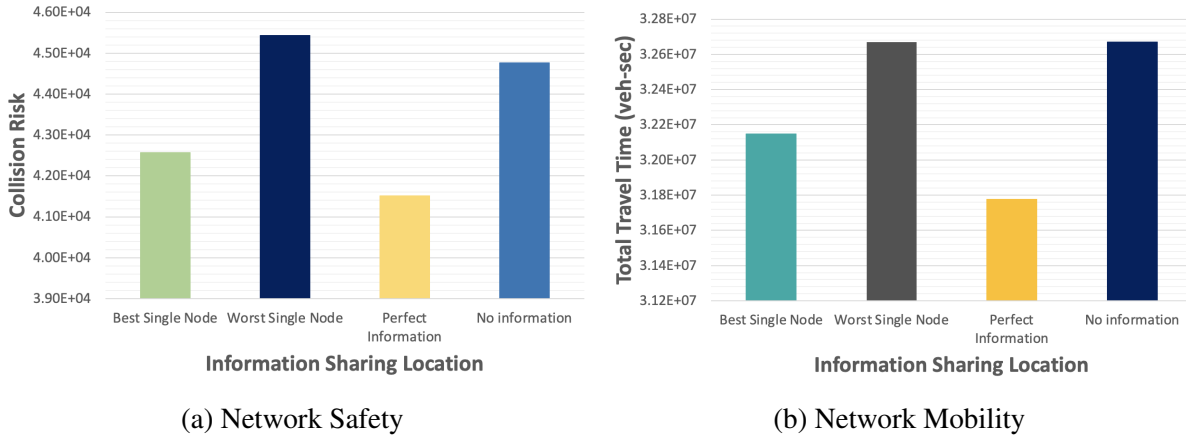
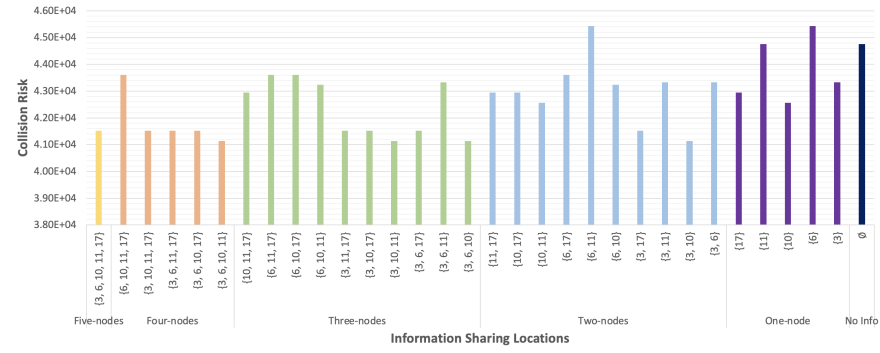


Figure 2.8: Impact of Perfect and no Information Sharing

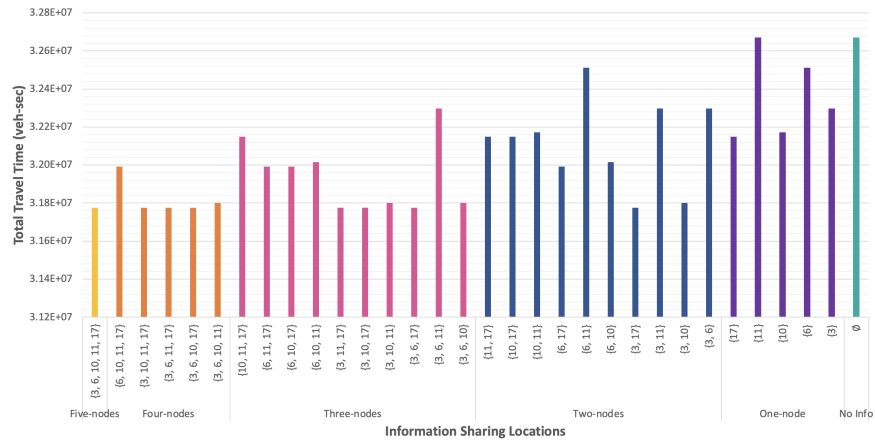
<sup>3</sup>. Notice that this is even better than the perfect information sharing strategy as shown in Figure 2.8a. This observation indicates that more information is not necessarily better for the whole system, which is consistent with observations in existing studies on “information paradox”(Acemoglu et al., 2018; Lindsey et al., 2014; Liu and Liu, 2018; Unnikrishnan and Waller, 2009). In terms of mobility, Figure 2.8b shows that if the information is shared at nodes 3 and 17, it provides the best system mobility outcome. To determine the Pareto optimal information sharing strategies, Figure 2.9c shows the Pareto frontier when both safety and mobility are considered. Information sharing at nodes {3,17} and nodes {3,10} are two Pareto optimal information sharing strategies with the least amount of information sharing nodes.

---

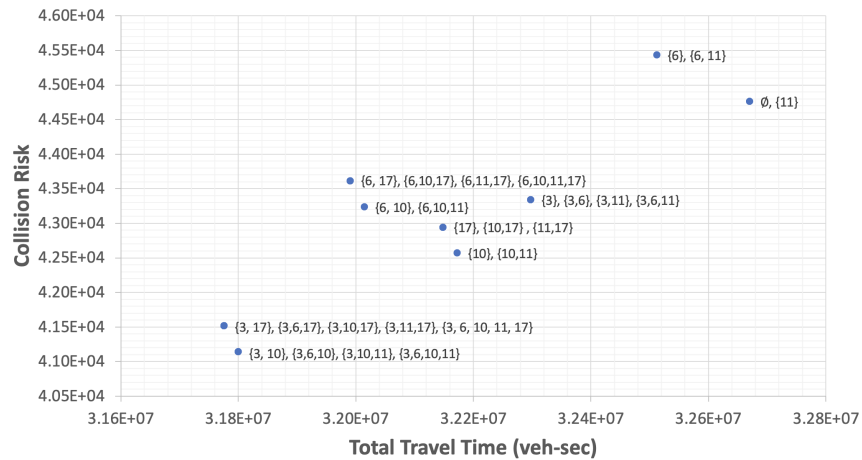
<sup>3</sup>When different combinations have the same impacts on the system, we prefer the combination with the least information sharing nodes to save installation costs.



(a) Network Safety



(b) Network Mobility



(c) Pareto Graph

Figure 2.9: Impact of at Information Sharing at Selected Nodes

### 2.4.2.2 Sensitivity of O-D Demand

In this section, we investigate the sensitivity of different levels of O-D demand on the average collision risk and the average travel time<sup>4</sup> of the network for both information shared and no information shared cases. We consider a range of O-D demand from 40% decrease to 40% increase from the base case O-D demand, with the results shown in Figure 2.10. From Figure 3.11b, we can see that a higher level of O-D demand results in a monotone increase in the system average travel time because of the non-linear congestion effects. The average travel time in the network is slightly higher when information is not shared than when information is shared for all O-D demand levels (Figure 3.11b). However, from the perspective of network safety, we observed a more complex pattern. As shown in Figure 2.10a, the average collision risk in the network decrease slightly when O-D demand increase from 60% to 80% and increase significantly from 80% to 120% when information is shared. The increase rate slows down when OD-demand increases from 120% to 140%. When information is not shared, the average collision risk is much higher than when information is shared, except when O-D demand is at 60% of the base case value. This is because information sharing affects system collision risk in two major ways. On one hand, the travelers receive information can make route adjustments to avoid experiencing high collision risks on the incident links. On the other hand, information sharing may encourage travelers to travel to the information nodes, which may have negative impacts on system collision risk. When the transportation network is not congested, collision risks may not be high on the incident links. But when information is shared at a node, travelers may still try to select the routes to access information updates and reroute afterward to maximize their own expected utility, which creates more negative externalities to the system compared with the no-information-sharing case.

---

<sup>4</sup>The reason we measure mobility and safety using average risk/time instead of total risk/time is to normalize the impacts by O-D demand level.

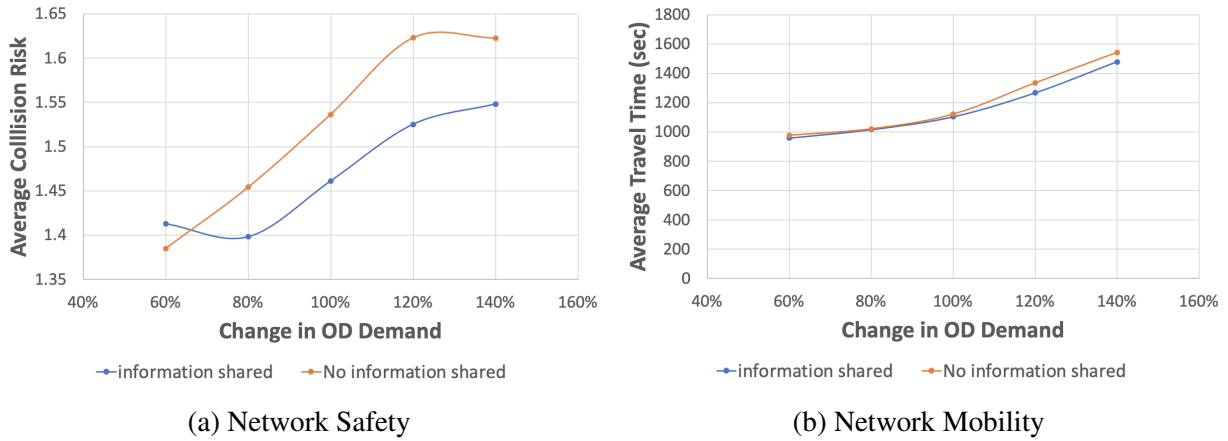


Figure 2.10: Effect of Change in O-D Demand

### 2.4.2.3 Sensitivity Analysis on Incident Severity

Figure 2.11 depicts the impact of various incident severity levels on the safety and mobility of the network for both information shared and no information shared cases. Low severity (20% capacity reduction) and high severity (80% capacity reduction) levels have been used and compared with the base case (50% capacity reduction). Figure 2.11 shows that both the network collision risks and the total travel time are slightly higher when no information is shared compared with when information is shared, which means information sharing is beneficial regardless of different incident severity levels. Figure 2.11b shows that the total travel time of the network increases as the incident severity level increases, which is intuitive since the overall network capacity decreases with a higher level of incident severity. This observation is consistent regardless of the information shared or not. However, for network safety, the relationship between the network collision risk and the severity of the incident may not be monotone especially when information is shared. For example, a more severe incident from the “Low” case to the “Base” case leads to a much lower level of network collision risks. This is because different levels of incident severity will increase

the travel time differently on those impacted links, which leads to different rerouting strategies of individual travelers when information is shared, resulting in different safety externalities on the system.

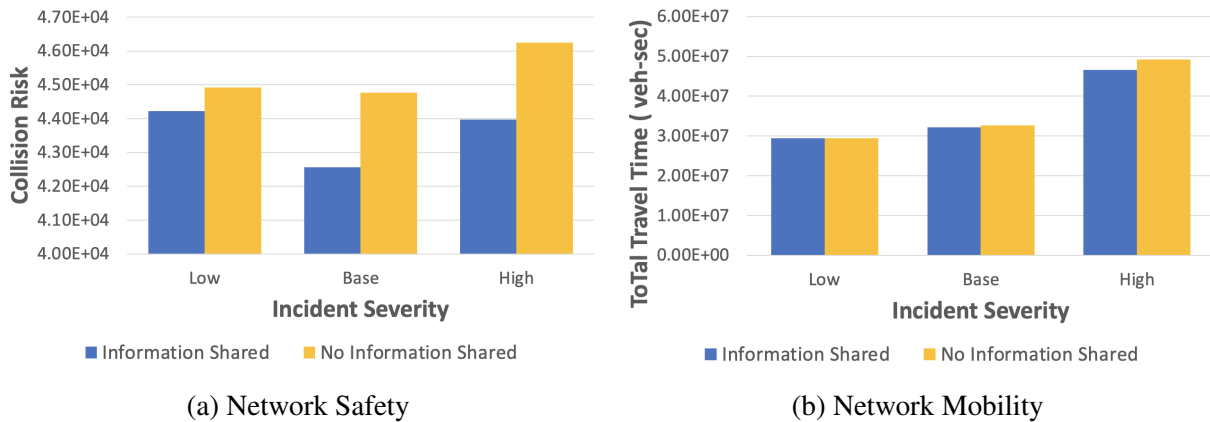


Figure 2.11: Effect of Incident Severity on Network Mobility and Safety

## 2.5 Discussion

In this study, we investigate the impact of information sharing locations on transportation network safety and mobility. We propose a transportation network modeling framework to model the adaptive routing behavior of CAVs given information updates at different information sharing locations. We estimate the traffic mobility parameters using traffic data for the Orlando network and the crash risk of each link from real crash data. The proposed methodology allows us to identify the Pareto optimal locations to share information with CAVs that helps to promote the safety and mobility of the whole network. Through numerical experiments, we found that: (1) the optimal information sharing strategies depend on specific network configurations and more/less information is not always better/worse for the network mobility and safety; (2) locational information sharing will encourage the traffic to travel through information nodes so that they can make informed

rerouting decision; (3) while mobility is monotone decreasing with the levels of OD demand and incident severity, network collision risks may have a more complex relationship, which can only be quantified with a network modeling perspective and real historical crash data.

This research can be extended in several directions. First, we assume there is only one universal piece of information in the network. How to extend the proposed methodology to consider heterogeneous information sharing remains to be solved. But extending the proposed SUER from two-stage to multi-stage will not change the overall modeling strategies. Second, in addition to information sharing locations, other aspects of information sharing strategies can be investigated, such as what information to share, to which group of CAVs to share information. Third, the study can be extended for mixed traffic that forms of CAVs and regular vehicles. Fourth, given increasing concerns of cybersecurity, investigating how transportation network safety and mobility will be influenced by erroneous information is also a valuable next step.

# **CHAPTER 3: OPTIMAL SPEED LIMIT CONTROL FOR NETWORK MOBILITY AND SAFETY: A TWIN-DELAYED DEEP DETERMINISTIC POLICY GRADIENT APPROACH**

## **3.1 Introduction**

Variable speed limit control (VSLC) has emerged as a promising approach for enhancing traffic safety, improving mobility, and reducing congestion in transportation systems. By dynamically adjusting speed limits based on real-time traffic and weather conditions, VSLC can reduce speed variance and increase average headway, leading to improved safety and smoother traffic flow (Lee et al., 2004; Park et al., 2008). It is a cost-effective measure that leverages existing infrastructure to alleviate both recurrent (Emmerink et al., 1995) and non-recurring congestion (Chen et al., 2016; Kwon et al., 2006), minimizing the need for infrastructure upgrades (Farrag et al., 2020; Hoogendoorn et al., 2013; Li et al., 2014). Furthermore, VSLC has the potential to reduce vehicle emissions (Bel and Rosell, 2013; Grumert et al., 2015) and dampen traffic shock waves (Hegyri et al., 2005b, 2008). Despite its benefits, optimizing VSLC strategies at a network level is challenging due to adaptive decision making and the interconnected nature of transportation networks. The impacts of VSLC placement on network safety and mobility need to be assessed, considering indirect effects such as changes in drivers' routing decisions and the potential for increased lane changing and local bottlenecks. Existing research primarily focuses on improving performance at a local intersection or link level, overlooking the broader impacts of VSLC on the entire network and considering the effects of driver rerouting. Recent studies using reinforcement learning approaches have shown promising results, but they often neglect the potential impacts of VSLC on surrounding links and fail to consider the cascading effects of rerouting. Limited research exists on the safety and mobility implications of VSLC in interconnected transportation networks, hindering

a comprehensive analysis of its deployment impacts.

The main goal of this study in this chapter is to improve network mobility and safety using VSLC. To achieve this goal, we have the following objectives. First, we proposed a policy gradient-based reinforcement learning approach and compare the model's learning process with other deep reinforcement learning models. Second, we use different traffic control strategies using VSLC to assist with vehicles' routing decisions and examine the network performance for these approaches. Third, we assess traffic events such as bottleneck and lane change improvement using VSLC in the network. Finally, we conduct a sensitivity analysis to determine the suitable location to implement VSLC that improves network performance. The main contribution of this study is threefold.

- 1) We develop an effective reinforcement learning (RL) algorithm that can adaptively control the VSLC devices to optimize network mobility and safety.
- 2) We determine the best location to place VSLC that improves mobility and safety of the network considering the rerouting behavior of the vehicles.

## **3.2 Literature Review**

Variable speed limit (VSL) implementation is a widely used traffic control method that can improve safety and reduce congestion by dynamically adjusting speed limits based on traffic conditions. Several methods have been proposed for VSL implementation, including rule-based, model-based, and RL-based approaches.

Earlier studies on variable speed limit control predominantly involved the development of simple reactive rule-based logic. Rule-based methods rely on predefined rules to determine the appropriate speed limit based on traffic conditions. Several studies have been conducted to improve traffic



mobility (Li and Ranjitkar, 2015; Lin et al., 2004; Papageorgiou et al., 2008) and safety (Abdel-Aty et al., 2006; Allaby et al., 2007; Lee et al., 2003, 2004) using rule-based approaches. Papageorgiou et al. (2008) proposed a flow-speed threshold-based VSL control algorithm and explored the impact of VSLs on traffic flow efficiency on motorways using real traffic data. The findings suggest that VSLs can improve flow efficiency by decreasing the slope of the flow-occupancy diagram and enabling higher flows at the same occupancy values in overcritical conditions. Li and Ranjitkar (2015) proposed a fuzzy logic-based algorithm to improve the effectiveness of VSL systems, which overcomes the sensitivity issues caused by crisp discretization in existing rule-based algorithms. The simulation results showed that the proposed algorithm outperforms the existing VSL algorithm by increasing the critical bottleneck capacity and reducing total travel time. Rule-based systems are relatively simple and require minimal computational resources, however, they may not be able to adapt to changing traffic patterns or handle complex traffic scenarios.

In contrast, model-based methods use mathematical models of traffic flow to predict the effects of different speed limits on traffic conditions, such as, model predictive control (MPC) (Frejo and Camacho, 2012; Han et al., 2017; Hegyi et al., 2005a, 2002), and feedback control approaches (Carlson et al., 2011, 2014; Iordanidou et al., 2014; Jin and Jin, 2015; Müller et al., 2015). MPC is a model-based control approach that uses a mathematical model of the system to predict its future behavior and optimize a control input over a finite prediction horizon. For instance, Han et al. (2017) proposed a novel approach using MPC for variable speed limit coordination to resolve freeway jam waves. Their proposed approach addresses the computational burden of existing MPC methods and resolves jam waves using a more accurate discrete first-order model. Simulation results show that the proposed MPC approach can resolve jam waves with real-time computation speed. One disadvantage of MPC is that it requires a mathematical model of the system, which can be difficult and time-consuming to develop and maintain. MPC also involves solving a complex optimization problem at each control interval, which can be computationally expensive.

On the other hand, in feedback control approaches, the controller sets the speed limits in order to maintain a system variable, e.g., traffic density, close to a defined critical/reference value. For example, in a study done by Jin and Jin (2015), the speed limit on the upper stream section of a lane-drop bottleneck is adjusted to maintain the traffic density at the bottleneck close to its critical capacity. One drawback of feedback control is that it may not be able to handle complex constraints and uncertainties as effectively as MPC. Feedback control may also lead to instability if the control gains are not properly tuned. Overall, model-based methods have advantages over rule-based approaches to variable speed limit control, which rely on fixed rules or heuristics. Model-based methods can adapt to changes in traffic conditions and optimize the speed limit for the current situation, whereas rule-based approaches may not be able to account for all possible scenarios. However, model-based methods can be computationally expensive, and the accuracy of the models depends on the quality of the input data.

RL-based approaches have been proposed in recent years to overcome the limitations of the model-based approaches. RL is a machine learning (ML) approach readily used in different applications of Intelligent Transportation Systems such as traffic signal control (Gregurić et al., 2020), demand assignment in shared autonomous mobility (Guériau and Dusparic, 2018), etc. In terms of VSLCs, RL approaches can train a smart controller that automatically sets speed limits on segments of the road network with VSL capabilities. In every RL application, the problem needs to be modeled by defining system states, actions, and reward functions. In RL-VSLCs, the actions are the speed limits and states can be different traffic parameters e.g., traffic density, average speed, current speed limit, etc. Additionally, the reward function should be defined to achieve the main goal of the controller which can be mobility or safety measurement.

The proposed RL-VSLCs in the literature mainly differ in terms of the selection of system states, reward functions, and the learning algorithms (Kušić et al., 2020b). Q-learning (QL) based approaches are prominent in the proposed RL-VSLCs (Li et al., 2017; Vrbanić et al., 2021; Walraven

et al., 2016). In Li et al. (2017) a QL-based VSL is proposed to reduce traffic congestion in freeway bottlenecks with total travel time as the reward function. Similarly Vrbanić et al. (2021); Walraven et al. (2016) have considered travel time-dependant rewards to alleviate congestion in transportation networks where they used the classic QL training algorithms. However, for a higher number of state-action pairs the QL approach becomes computationally intractable known as the curse of dimensionality. Additionally, in QL approaches, the state space should be discrete space which is limiting for real-world applications. Researchers have proposed W-learning (Kušić et al., 2020a), decentralized multi-agent QL approaches (El-Tantawy et al., 2013; Wang et al., 2019a) to overcome the curse of dimensionality of the classic QL method. Other studies have also proposed the k-NN method to map the state-action pair into lower dimensions to help with the computation efficiency (Li et al., 2017; Schmidt-Dumont and Van Vuuren, 2019). Recently, several RL approaches are proposed for controlling the VSLCs which are both computationally efficient and can take continuous state spaces into account (Gregurić et al., 2022; Wu et al., 2020c). For example, authors in Wu et al. (2020c) proposed an actor-critique algorithm for differential variable speed limits (DVSL) control, which can improve traffic conditions, and safety, and reduce emissions. The proposed model is tested on a simulated freeway recurrent bottleneck, and the results show improved efficiency, safety, and reduced emissions.

The main limitation of the mentioned studies is that they have only considered a segment of a road or a highway bottleneck to enforce VSLCs for controlling traffic. Zhu and Ukkusuri (2014) proposed an R-Markov Average Reward Technique reinforcement learning algorithm (RMART) to obtain the optimal speed limit scheme in the Sioux fall network, which shows a reduction of 18% and 20% in total travel time and CO emissions, respectively, compared to the base case of the non-speed limit control. However, they haven't investigated how VSLCs can impact the network's overall traffic and how they might influence the routing choice of drivers.

### 3.3 Problem Statement

We use the four-node network Figure 3.1, which is a one-way network with vehicles having two different paths to reach their destination, to demonstrate our problem statement. For this study, vehicles can receive information about the road (i.e., real travel time) to determine the shortest path to reach the destination. Since vehicles are allowed to choose their path based on real-travel time on each path which is subjected to change by the admission of a new vehicle in the path, vehicles need to decide to reroute to reach their destination. This decision-making process creates a bottleneck before reaching the routing links (as shown in Figure 3.1) and leads to traffic oscillation in the routing links (as shown in Figure 3.2). Traffic oscillation occurs because vehicles prioritize one route over the other, resulting in an alternating peak of occupancy for each route. For example, when the occupancy is maximum in Link 2-3 (Figure 3.2a), it is minimum in Link 2-4 (Figure 3.2b), indicating that most vehicles take the route 1-2-3-4-6 at that point. We want to control the speed limit of one of the routing links in the congested network and improve the safety and mobility of the entire network. A routing link is a specific segment or section of a route that connects an origin to a destination and is a part of one or more potential routes. This entails determining the rewards for the VSLC to provide the speed limit that helps enhance network mobility and safety. Additionally, we want to dampen the traffic oscillation in the routes via VSLC implementation and observe network mobility and safety. The simple network used in this study is in fact very crucial as we intend to understand the vehicles' routing decisions based on real-time information which can be harder to interpret from a larger network. The results from this study can be further used to implement VSLC in larger networks.

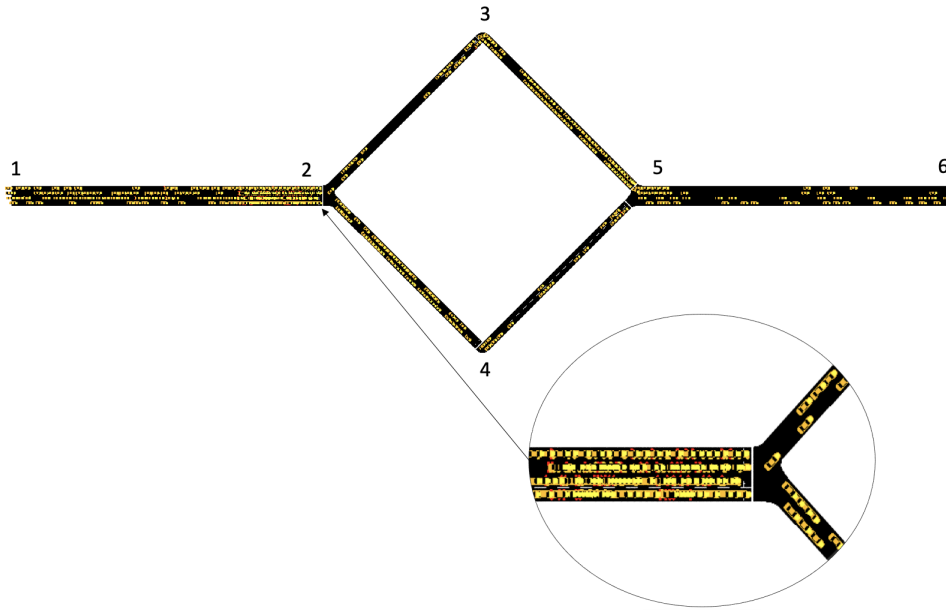
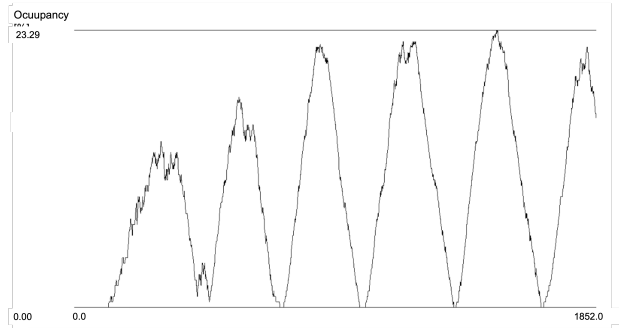


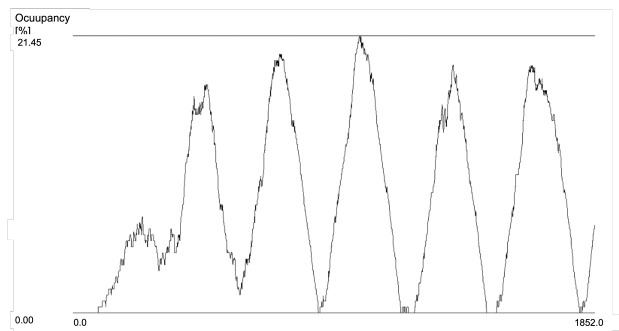
Figure 3.1: Problem Setting

### 3.4 Theoretical Background

This section presents the theoretical background of the RL approach. This is because to improve the safety and mobility of the network by placing VSLC, the RL approach is the state-of-the-art approach. Before we discuss the proposed model that we used for this study, we first offer a brief discussion of RL, based on which we describe how RL is enhanced by integrating deep learning, which gives birth to DRL. Followed by that, we discuss some of the latest algorithms capable of handling problems with continuous action space.



(a) *Link 2 – 3*



(b) *Link 2 – 4*

Figure 3.2: Traffic oscillation in the routing links

### 3.4.1 Basic Overview of Reinforcement Learning

RL is a form of ML approach, where the goal of an agent is to interact with an environment to learn and optimize its behavior from its interaction experience. In other words, RL is a sequential decision process where an agent needs to make a decision at time  $t$  after interacting with the environment. RL algorithm can be model-based or model-free. Having a full model of the environment means knowing all the state transition probabilities with certainty. Contrary to that, model-free algorithms, such as Q-learning, and deep Q-learning (DQL), rely on the Markov decision process (MDP) to keep track of the decision process in the environment. In our study with the four-node

network, at a time  $t$ , an agent experience the state  $s_t$  (i.e., occupancy of vehicles in each link) and takes action  $a_t$  (i.e., implement a speed limit in link/links). This state and action help achieves reward  $r_t$  and transforms the environment into a new state  $s_{t+1}$  at time  $t + 1$ . The transition process MDP can be represented by  $p(s_{t+1}|s_t, a_t)$  where the transition to a new state depends on the immediately previous state and action, not any other states and actions (Markov property). The interaction between the agent and the four-node network environment is shown in Figure 3.3.

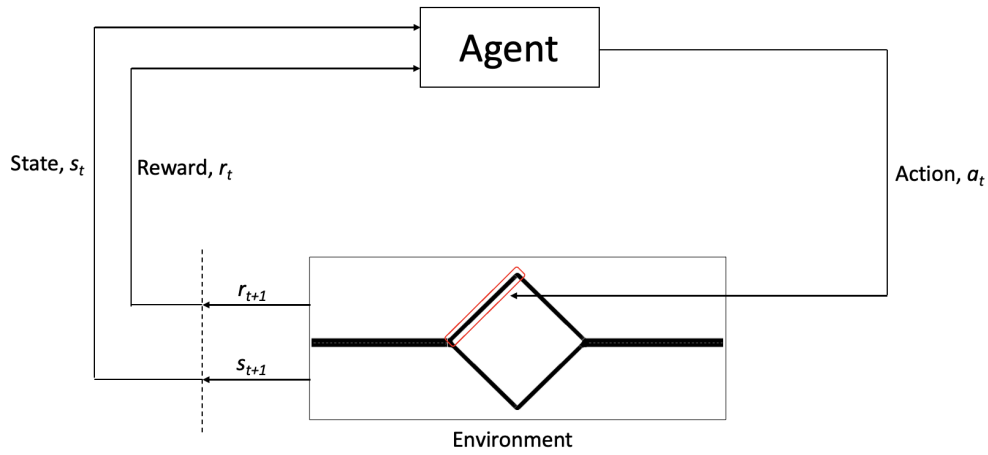


Figure 3.3: Interaction between agent and environment

The goal of an agent is to optimize the cumulative reward, which is represented as the "return" over the entire time period. At a time  $t$  return,  $G_t$  is the total accumulated reward received by an agent over the time  $T$ ,

$$G_t = r_{t+1} + r_{t+2} + r_{t+3} + r_{t+4} + \dots + r_{t+1+T} \quad (3.1)$$

Instead of using Equation (3.1), reward over time is discounted using a discount rate  $\gamma$ . Discounting will make the agent care about the immediate reward more than the future reward as the future reward will be heavily discounted. This return is shown in the following equation.

$$G_t = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \gamma^3 r_{t+4} + \dots + \gamma^T r_{t+1+T} = \sum_{k=0}^T \gamma^k r_{t+1+k} \quad (3.2)$$

The returns at successive time steps are related to each other which is shown below.

$$G_t = r_{t+1} + \gamma G_{t+1} \quad (3.3)$$

The "expected return" is the average return that an agent can expect to receive by following a given policy in a given environment. It is calculated by taking the expected value of the return over all possible sequences of actions and rewards, weighted by the probability of each sequence occurring. For brevity in presentation, we discard the probability which gives us the same expression as the "return". From here forward, the return expression is used to present the expected return. Depending on the approach to achieve the optimal goal (optimizing expected return), the model-free algorithm has been categorized into three groups: (1) value-based algorithms, (2) policy-based algorithms, and (3) actor-critic algorithms. Value-based algorithms are a type of RL algorithm that estimates the value function of a state or state-action pair in order to determine the optimal policy. The most popular value-based algorithm is Q-learning, which uses the Bellman equation to iteratively update the Q-value of a state-action pair based on the expected reward of the next state.

The Q-value of a state-action pair  $(s_t, a_t)$  is defined as the expected cumulative reward for taking action  $a_t$  in state  $s_t$  at time  $t$  and following the optimal policy thereafter. The Q-value is defined recursively as:

$$Q(s_t, a_t) = \mathbb{E}[r(s_t, a_t) + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) | s_t, a_t] \quad (3.4)$$

where  $r(s_t, a_t)$  is the immediate reward received at time  $t$  after taking action  $a_t$  in state  $s_t$ ,  $s_{t+1}$



is the resulting state after taking action  $a_t$  in state  $s_t$ ,  $\gamma$  is the discount factor that determines the importance of future rewards, and  $\max_{a_{t+1}} Q(s_{t+1}, a_{t+1})$  is the maximum Q-value for all possible actions  $a_{t+1}$  in state  $s_{t+1}$ . The optimal policy is obtained by selecting the action that maximizes the Q-value for each state:

$$\pi^*(s) = \underset{a}{\operatorname{arg\,max}} Q(s, a) \quad (3.5)$$

In Q-learning, the Q-values are updated iteratively using the following update rule at each time step  $t$ :

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [r(s_t, a_t) + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)] \quad (3.6)$$

where  $\alpha$  is the learning rate that determines the step size of the update. Q-learning converges to the optimal Q-values and optimal policy under certain conditions, such as the Markov property and a sufficiently small learning rate. However, it may suffer from overestimation or underestimation of Q-values and may not be suitable for large state spaces or continuous action spaces.

In contrast, policy-based algorithms are a class of RL algorithms that learn a parameterized policy directly, without computing the value function. The policy is typically represented as a function that maps states to actions, and the goal is to optimize this function to maximize the expected return. The optimization problem for a policy-based algorithm can be stated as follows:

$$J(\theta) = \mathbb{E}_{\mu_\theta} \left[ \sum_{t=0}^{T-1} G_t \right] \quad (3.7)$$

Here  $J(\theta)$  is the objective function that we want to maximize,  $\theta$  are the parameters of the policy,  $\mathbb{E}_{\mu_\theta}$  denotes the expected value under the policy  $\mu$  parameterized by  $\theta$ . To optimize the policy, the gradient of the objective function with respect to the policy parameters is computed, which can be

done using the policy gradient theorem:

$$\nabla J(\theta) = \mathbb{E}_{\mu_\theta} \left[ \sum_{t=0}^{T-1} \nabla_{\theta} \log \mu_{\theta}(s_t, a_t) Q(s_t, a_t; \theta) \right] \quad (3.8)$$

where  $\nabla J(\theta)$  is the gradient of the objective function,  $\nabla_{\theta} \log \mu_{\theta}(s_t, a_t)$  is the derivative of the log probability of taking action  $a_t$  in state  $s_t$  at time  $t$  with respect to the policy parameters  $\theta$ .  $Q(s_t, a_t)$  is the state-action value function (same as Equation 3.4) under policy  $\mu$  parameterized by  $\theta$  at time  $t$ . The objective function  $J(\theta)$  is defined in terms of the expected return of the policy, which can be expressed in terms of the state-action value function  $Q(s_t, a_t)$ . This is because the expected return of the policy is equal to the sum of the expected rewards obtained from the current state  $s_t$  and action  $a_t$ , plus the expected return from the next state  $s_{t+1}$ . The update rule for the policy parameters can then be derived by taking a step in the direction of the gradient:

$$\theta \leftarrow \theta + \alpha \nabla J(\theta) \quad (3.9)$$

where  $\alpha$  is the learning rate. In practice, the policy gradient is estimated using Monte Carlo sampling or temporal difference learning. Monte Carlo sampling involves running a complete episode and computing the return for each state-action pair, while temporal difference learning involves updating the policy parameters after each step based on the observed reward and the predicted value of the next state. Overall, policy-based algorithms have several advantages over value-based algorithms, such as the ability to handle continuous action spaces and the ability to learn stochastic policies. However, they may be slower to converge and can suffer from high variance in the gradient estimates.

Both value and policy-based algorithms have their own cons. For example, value-based algorithms cannot handle continuous and high-dimensional action space, which policy-based algorithms can

handle. However, policy-based algorithms may have large variances (Konda and Tsitsiklis, 1999). Moreover, with changes to policy, the new gradient is estimated irrespective of the previous policies. Therefore, the agent is not learning with respect to the accumulation of previous information. To overcome these limitations, an actor-critic algorithm combines the benefits of both value-based and policy-based methods by using two separate neural networks: an actor and a critic. The actor network outputs an action based on the current state, while the critic network estimates the value of the state-action pair. During training, the critic evaluates the actions taken by the actor and provides feedback in the form of a temporal-difference (TD) error, which is the difference between the estimated value of the current state-action pair and the estimated value of the next state. The actor network then uses this feedback to update its policy and select better actions in the future.

### 3.4.2 Deep Reinforcement Learning

DRL algorithms, which use deep neural networks (DNNs) to approximate the value functions, have shown significant progress in solving complex problems. However, the use of neural networks for function approximation faced instability and convergence issues when dealing with larger and more complex problems (Tesauro et al., 1995). The primary cause of this instability was the sequential nature of the training dataset, which gave rise to auto-correlation issues and non-i.i.d. data distribution. To address these limitations, the DQN algorithm was introduced. DQN addressed the instability of DRL by employing two key measures: experience replay and target networks.

#### 3.4.2.1 *Experience Replay*

Experience replay is a widely used technique in DRL where the agent stores a buffer of past experiences in the form of state, action, reward, and next state tuples encountered during its interaction with the environment. The agent then randomly samples a mini-batch of experiences from this

buffer to update its neural network. Experience replay offers several benefits to DRL. Firstly, by randomly sampling experiences, temporal correlations in the sequence of experiences can be eliminated, which otherwise would lead to a biased estimate of the value function or the policy. Secondly, it helps to stabilize the learning process as the mini-batch used for learning is less likely to be influenced by noisy or outlier experiences. Thirdly, experience replay enables the agent to learn from diverse experiences, including those encountered in the distant past, which can be crucial in tasks requiring the agent to recall past information. Finally, experience replay can enhance data efficiency by enabling the agent to reuse past experiences, allowing it to learn from a small set of experiences multiple times instead of relying only on the most recent experience.

#### *3.4.2.2 Target network*

In DRL, the target network plays a crucial role in stabilizing the training process and improving the convergence of the learning algorithm. It is essentially a duplicate of the neural network used by the agent to estimate the value function or policy, but it is not updated as frequently as the primary network. Instead, it generates target values for the update of the primary network. Using a target network has several advantages. Firstly, it helps to improve the stability of the training process by reducing the correlation between the target values and the current estimates. This is because the target values are generated using a separate network that has not been updated for a few iterations, which reduces the impact of noisy updates on the primary network. Secondly, using a target network can reduce the oscillations in the learning process that can occur when the primary network is updated using estimates of the value function that are themselves changing rapidly during training. Thirdly, the target network can also help to improve the convergence of the learning algorithm by providing more stable target values for the update of the primary network. This leads to faster convergence and more consistent performance of the agent. Lastly, using a target network can also help to address the problem of overfitting, which can occur when

the primary network is over-optimized to the specific experiences encountered during training. By regularizing the learning process and preventing overfitting, a separate target network can help to achieve better generalization and robustness.

An early RL model that used both these features is DQN. DQN is a model-free RL algorithm that combines Q-learning with DNNs to learn a Q-value function that estimates the expected return for a given state-action pair. The Q-value function is represented by a DNN with weights  $\theta$ . Given a state  $s$ , the Q-value function approximates the expected discounted return for each action  $a$  as follows:

$$Q(s, a; \theta) \approx Q^*(s, a) \quad (3.10)$$

The DQN algorithm uses experience replay and a target network to improve the stability of the Q-value function learning process. Experience replay involves storing experience tuples  $(s, a, r, s')$  in a replay buffer, which is randomly sampled during training to break the temporal correlation of the data. The target network is a copy of the Q-network with frozen weights, which is used to compute the target Q-value for the next state in the TD error calculation. The parameters  $\theta$  of the Q-network are updated by minimizing the mean squared error (MSE) loss between the predicted Q-value and the target Q-value:

$$\mathcal{L}(\theta) = \mathbb{E}[(r + \gamma \max_{a'} Q'(s', a'; \theta') - Q(s, a; \theta))^2] \quad (3.11)$$

In Equation (3.11) instead of a second forward pass in the DNN to obtain  $Q'(s', a'; \theta')$ , a target network is used. At the beginning of training, the target network parameter  $\theta'$  is set to be the same as the original DNNs, i.e.,  $\theta$ . Unlike  $\theta$ ,  $\theta'$  is kept frozen for a certain number of time steps before an update. Then,  $\theta'$  is updated to the current value of the DNN parameter  $\theta$ . Then,  $\theta$  is updated

through backpropagation and gradient descent.

Although DQN and its variants have been used for decades to solve various problems (Chen et al., 2018; Zhang et al., 2019; Zhu et al., 2021), it is difficult to solve real-world issues using these approaches. This is because DQN is suitable for a few possible actions. Therefore environments with continuous action space would not perform well using DQN and its variants. DDPG is a newer algorithm that uses concepts from both policy gradient and DQN to enhance the capability of the algorithm to handle continuous action space.

DQN achieved significant success in playing Atari games, outperforming previous state-of-the-art algorithms and even matching human-level performance on some games. Since then, various extensions of DQN have been developed, such as Double DQN, Dueling DQN, and Rainbow, further improving the performance of DRL in various tasks.

### *3.4.2.3 Deep Deterministic Policy Gradient*

DDPG is an actor-critic algorithm that can handle continuous action space (Sutton and Barto, 1999). The goal of the actor is to decide which action to take and the critic evaluates the action and informs the actor how good the action is. Based on the actor-critic approach, DDPG has the actor network and the critic network. Similar to DQN, DDPG uses two subnetworks to represent both the actor and the critic network: (1) online network, and (2) target network. These DNNs are parameterized by  $\theta$ . Since there are two different types of networks, actor and critic networks, the subnetworks are parameterized by  $\theta^\mu$  and  $\theta^Q$  accordingly. DDPG and the actor-critic framework differ in their choice of policy representation. While actor-critic methods typically use stochastic policies and rely on policy gradient methods to update them (Sutton et al., 1999), DDPG employs a deterministic policy (denoted by  $\mu$ ) that is updated via the deterministic policy gradient. This eliminates the large variance faced in policy gradient methods, which can lead to slow convergence

and unstable training (Silver et al., 2014).

Similar to DQN, after assigning random weights ( $\theta^\mu$  and  $\theta^Q$ ) to the DNNs, DDPG uses experience replay where DNN is trained using a randomly selected experience (denoted as  $N$  sets). The actor network depends on the feedback from the critic network; therefore critic network activity takes place first in the training process. The critic network is updated using the gradient descent on the loss function

$$\nabla_{\theta^Q} \mathcal{L}(\theta^Q) = \mathbb{E}_{s,s',a,r \sim N} [\nabla_{\theta^Q} (y - Q(s,a;\theta^Q))^2] \quad (3.12)$$

where,

$$y = r + \gamma Q'(s', a'; \theta^{Q'}) \quad (3.13)$$

Equation (3.13) is the definition of the target Q-value, which is used in the update of the critic network in Equation 3.12. It is a weighted sum of the current reward ( $r$ ) and the estimated value of the next state-action pair, represented by  $Q'(s', \mu'(s'; \theta^{\mu'}); \theta^{Q'})$ . The discount factor  $\gamma$  determines the relative importance of immediate and future rewards. The target Q-value is computed using the target critic network  $Q'$ , which is a copy of the online critic network with frozen weights. The target actor network  $\mu'$  is used to determine the action  $a' = \mu'(s'; \theta^{\mu'})$  for the next state  $s'$ .

The actor network updates with the assistance of the critic network. The online actor network passes action ( $\mu(s)$ ) to the online critic network which helps achieve the action's gradient  $\nabla_a Q(s, \mu(s))$ . With its own optimizer, the parameter  $\theta^\mu$ 's gradient  $\nabla_{\theta^\mu} \mu$  can be derived. With these two gradients, the policy update gradient for the online actor network is as follows.

$$\nabla_{\theta^\mu} J = \mathbb{E}_{s \sim N} [\nabla_a Q(s, a; \theta^Q) \nabla_{\theta^\mu} \mu(s; \theta^\mu)] \quad (3.14)$$

DDPG utilizes a soft update mechanism for the target networks, which allows for a gradual change in the target network parameters. This is in contrast to hard updates, which cause abrupt changes to the target network and can lead to instability in the learning process. Soft updates can lead to more stable learning and reduce the risk of overfitting, as the target network's parameters change gradually over time. The soft target is shown as follows where  $\tau$  is less than 1.

$$\theta^{\mu'} \leftarrow \tau\theta^{\mu} + (1 - \tau)\theta^{\mu'} \quad (3.15)$$

$$\theta^{Q'} \leftarrow \tau\theta^{Q} + (1 - \tau)\theta^{Q'} \quad (3.16)$$

### 3.5 Proposed Algorithm

This section presents the algorithm chosen for the study. In recent years, policy-based algorithms such as DDPG (Lillicrap et al., 2015) gained popularity in different fields, including, robotics (Jesus et al., 2019; Xu et al., 2018), electricity supply industry (Liang et al., 2020), energy system (Li et al., 2021a; Wu et al., 2020a), wireless communication (Qiu et al., 2019), transportation (Casas, 2017; Wang et al., 2019b; Wu et al., 2020b) and even on VSL related studies (Gregurić et al., 2022; Wu et al., 2020c).

Although DDPG is a relatively new algorithm and has provided promising results for real-world problems in different research fields, Fujimoto et al. (2018) showed that DDPG suffers from the following issues:

- 1) **Overestimation Bias:** The overestimation bias in DDPG algorithm occurs when the learned value function overestimates the true value of the state-action pairs. This overestimation bias can be problematic as it can lead to unstable and suboptimal behavior of the algorithm.



The overestimation bias arises when the policy and value networks are jointly learned using the same samples. The policy improvement step in DDPG updates the policy based on the current value estimate, which can be inaccurate due to the overestimation bias. This inaccurate value estimate can then lead to the selection of suboptimal actions by the policy, which in turn can lead to poor performance of the algorithm.

- 2) Accumulation Error: The accumulation error in DDPG occurs because the value function estimate is based on an estimate of a subsequent state, leading to a buildup of error over time. This error is exacerbated in a function approximation setting resulting in residual TD-error after each update. This can lead to large overestimation bias and suboptimal policy updates.

In this section, we propose the Twin-delayed deep deterministic policy gradient (TD3) approach (originally proposed by Fujimoto et al. (2018)) to implement VSLC for improving network mobility and safety. Figure 3.4 shows the structure of TD3 showing different elements of TD3. The different parts of TD3 and its implementation are discussed in the remaining section.

### 3.5.1 Deep Network Architecture

TD3 consists of two independent critic networks and one actor network, each having an online and a target network. The actor or policy network receives the current state of the simulation and outputs exact actions, instead of a probability distribution over actions; this, in turn, contributes to the fact that TD3 is "deterministic". The actor network consists of three fully connected layers with rectified linear activation functions (ReLU). The first two layers have 200 and 100 neurons, respectively, and use ReLU activation functions. The final layer has 50 neurons and does not use an activation function. The output of this layer is the action which is a continuous value

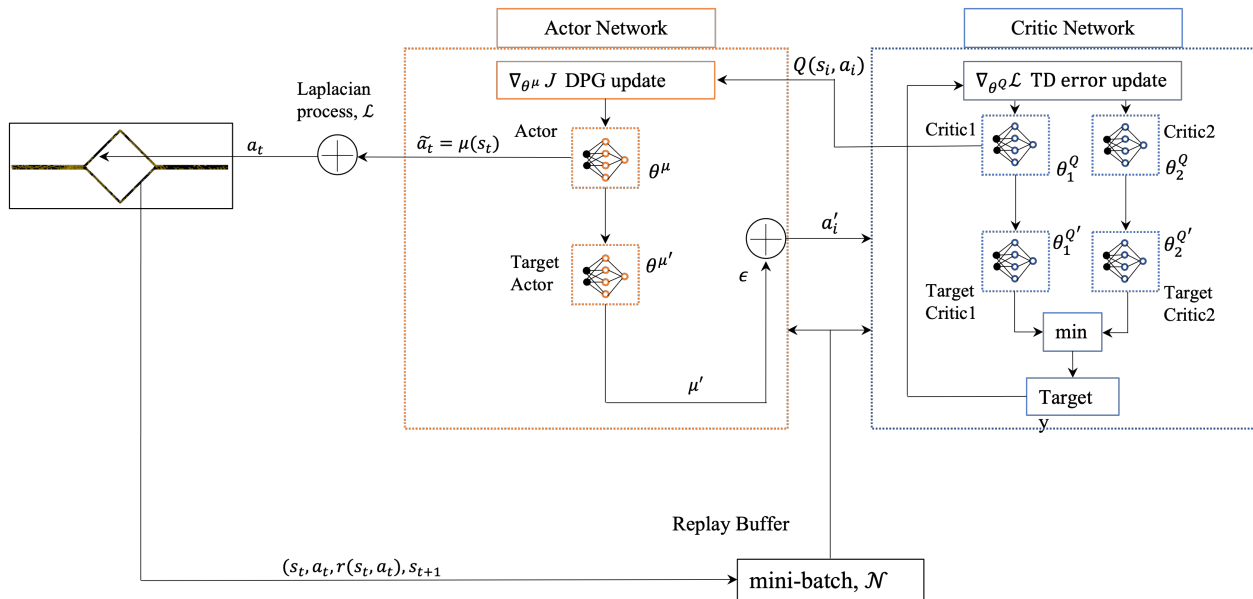


Figure 3.4: Structure of TD3

representing the speed limits for VSLCs. Later we show that we add noise in the actions to provide randomness to assist in exploring different action spaces. When interfacing with Simulation of Urban MObility (SUMO), it is necessary to convert the speed limit action output by the actor network to the appropriate unit used by SUMO, which is meters per second (m/s) (See details in Section 3.5.5).

The critic network receives the current state of the simulation plus the action generated by the actor and outputs the Q-values associated with them. TD3 uses a clipped variant of Double Q-learning where two independent critic networks are used to estimate two Q-values for each state-action pair. Each critic network has several fully connected layers with ReLU. Over the course of these layers, the neurons of each layer decrease as the number of layers increases (200,100, 50, and 1 neurons respectively).

The critic network in TD3 takes as input the current state of the simulation and the action generated

by the actor, and produces Q-values associated with them. The architecture of each critic network comprises multiple fully connected layers with ReLU. The number of neurons in each layer progressively decreases as the number of layers increases, with the final layer containing only one neuron. Specifically, the critic network consists of four layers with 200, 100, 50, and 1 neurons, respectively.

### 3.5.2 Online Update mechanism

Online network updates for critic and policy/actor networks are different for TD3 than for DDPG which is described as follows.

#### 3.5.2.1 *Online Critic Network Update*

To mitigate overestimation bias in Q-learning, TD3 employs a clipped variant of Double Q-learning that utilizes two independent critic networks to estimate two Q-values for each state-action pair. During the update process, the minimum of the two Q-value estimates is used as the target value in the Bellman equation update, thereby reducing overestimation bias (Equation 3.17). However, when the difference between the two Q-value estimates exceeds the clipping range, the target value is clipped to lie within this range.

To further enhance the stability of the learning process, TD3 adopts a technique called target policy smoothing when forming the target critic. This involves introducing noise to the target action generated by the target actor network during the Bellman equation update. The amount of noise added is determined by a small standard deviation clipped Gaussian distribution (Equation 3.18). The resulting smoothed action is then used as the target action for computing the Q-values in the critic network update. By regularizing the updates and preventing overfitting to the current policy,

target policy smoothing reduces instability in the learning process.

$$y_i = r_i + \gamma \min_{j=1,2} Q'_j(s_{i+1}, (\mu'(s_{i+1}; \theta^{\mu'}) + \varepsilon); \theta_j^Q) \quad (3.17)$$

$$\varepsilon \sim \text{clip}(\mathcal{N}(0, \sigma), -c, c) \quad (3.18)$$

To update the online critic network, the mean-squared error between the predicted Q-value and the target Q-value is calculated and used to compute the gradient with respect to the critic network parameters (Equation 3.19 and 3.20). The weights of the online critic networks ( $\theta_1^Q$  and  $\theta_2^Q$ ) are then updated using an optimizer, such as stochastic gradient descent (SGD) or Adam, to minimize the mean-squared error.

$$\nabla_{\theta_1^Q} \mathcal{L}(\theta_1^Q) = \frac{1}{N} \sum_i [\nabla_{\theta_1^Q} (y_i - Q_1(s_i, a_i; \theta_1^Q))^2] \quad (3.19)$$

$$\nabla_{\theta_2^Q} \mathcal{L}(\theta_2^Q) = \frac{1}{N} \sum_i [\nabla_{\theta_2^Q} (y_i - Q_2(s_i, a_i; \theta_2^Q))^2] \quad (3.20)$$

### 3.5.2.2 Online Actor Network Update

In TD3, the actor network is updated using the deterministic policy gradient (DPG) algorithm. The gradient of the expected return with respect to the actor parameters is computed using the chain rule, and this gradient is used to update the actor weights in the direction that increases the expected return.

$$\nabla_{\theta^\mu} J = \frac{1}{N} \sum_i [\nabla_a Q_1(s, a; \theta_1^Q)|_{s=s_i, a=\mu(s_i)} \nabla_{\theta^\mu} \mu(s; \theta^\mu)|_{s=s_i}] \quad (3.21)$$

TD3 uses delayed policy updates, which means that the actor network is updated less frequently than the critic network, as opposed to every critic update in DDPG. This further improves the stability of the learning process

### 3.5.3 Target Update mechanism

Similar to DDPG, the learning process is stabilized using a "soft update". This is accomplished by gradually changing the target network parameters. In the "hard update" the target weights are abruptly updated based on the online networks. The soft update is done at regular intervals using a copy of actor and critic networks by slowly maintaining the learned networks:

$$\theta^{\mu'} \leftarrow \tau\theta^{\mu} + (1 - \tau)\theta^{\mu'} \quad (3.22)$$

$$\theta_1^{Q'} \leftarrow \tau\theta_1^Q + (1 - \tau)\theta_1^{Q'} \quad (3.23)$$

$$\theta_2^{Q'} \leftarrow \tau\theta_2^Q + (1 - \tau)\theta_2^{Q'} \quad (3.24)$$

Equation 3.22, Equation 3.23, and Equation 3.24 are the soft updates for target actor, and the two target critic networks respectively. Here in this study,  $\tau$  is 0.005.

### 3.5.4 Exploration method

For our proposed model, it is essential for the model to learn as much behavior as possible. Unlike discrete action space, we cannot choose an action randomly in our environment and explore all the actions. The continuous action space does not have this advantage. The advantage off-policy

algorithms such as TD3, suitable for models with continuous action is that the exploration can be separate from the learning algorithm. We add noise to randomize the action using a Laplacian process (Jinnai et al., 2019) which follows the Laplace distribution, also known as a double exponential distribution as shown in Equation (3.25)

$$\mathcal{L}(\tilde{a}_t|b_t) \sim \frac{1}{2b_t} \exp\left(-\frac{\tilde{a}_t}{b_t}\right) \quad (3.25)$$

Here,  $\tilde{a}$  is the action we get from the model and  $b_t$  is the scale parameter referred to as diversity that decays with each learning period. By using Equation (3.25), we get a random action. This action is further clipped using Equation 3.26 at a time  $t$  and fed into SUMO simulation.

$$a_t = clip(\mathcal{L}(\tilde{a}_t|b_t), -c, c) \quad (3.26)$$

### 3.5.5 State, Reward, and Action for the TD3 model

In order to design an effective RL model, it is crucial to define appropriate states, actions, and rewards. A thorough literature review was conducted to identify the commonly used states and rewards in previous studies related to improving safety and mobility using variable speed limits and lane control. Occupancy and density were frequently used as states (Li et al., 2017; Wu et al., 2020c; Zhu and Ukkusuri, 2014) while few studies used speed (Kušić et al., 2020a; Walraven et al., 2016) or speed and density pair (Wang et al., 2019a) to define the state. The most common reward function for improving mobility is the total time spent (*TTS*) or total travel time (*TTT*) (Kušić et al., 2020a; Walraven et al., 2016; Wu et al., 2020c; Zhu and Ukkusuri, 2014). Time-to-collision (*TTC*) was commonly used as the reward function for improving safety (Peng and Xu, 2021; Wang et al., 2019a), while emergency braking vehicle was used as the reward for a study by Wu et al. (Wu et al., 2020c). The state, reward, and action for the proposed TD3 model are discussed below.

### 3.5.5.1 State:

The aforementioned studies mostly used occupancy and in some studies density as the state for their models. For the TD3 model, we have chosen occupancy of all edges as the state. Occupancy is the ratio of the sum of the lengths of the vehicles to the length of the road section in which those vehicles are present, usually represented as a percentage. The motivation is based on the observation of the occupancy of each edge, the TD3 algorithm will determine the appropriate actions for mobility and safety.

### 3.5.5.2 Reward:

rewards for different objectives

#### Objective - Improving safety:

*TTC* is a commonly used metric in the field of vehicle safety to assess the likelihood of a collision between two vehicles. It is defined as the estimated time it would take for two vehicles to collide if they maintain their current trajectory and speed (Hydén, 1996). The formula is as follows.

$$TTC_i = \begin{cases} \frac{X_i(t) - X_{i-1}(t) - l_i}{\dot{X}_i(t) - \dot{X}_{i-1}(t)}, & \text{if } \dot{X}_i(t) > \dot{X}_{i-1}(t) \\ \infty, & \text{Otherwise} \end{cases} \quad (3.27)$$

where:

$\dot{X}_i$  : the speed of the vehicle  $i$

$X_i$  : the rear position of the vehicle  $i$

$l_i$  : the length of vehicle  $i$

$i - 1$  : the vehicle in front of vehicle  $i$

The denominator is the relative velocity between the two objects. The relative velocity is calculated as the difference between the velocity of the two objects. If the two objects are moving in the same direction, the relative velocity is the difference between their velocities. If the two objects are moving in opposite directions, the relative velocity is the sum of their velocities.

$TTC$  is a useful metric for safety reward in RL because it provides a direct measure of the safety of a given action. In particular, if an agent takes an action that leads to a decrease in  $TTC$ , then that action can be considered unsafe and penalized accordingly. On the other hand, if an agent takes an action that leads to an increase in  $TTC$ , then that action can be considered safe and rewarded accordingly. Here, we propose a new definition of  $TTC$  which is the logarithm (Figure 3.5) of  $TTC$  as shown below.

$$TTC_{i,new} = \log(TTC_i) \quad (3.28)$$

There are certain advantages to using Equation(3.28) as the safety reward function for a RL model. First, by taking the logarithm of the time-to-collision, the range of values for  $TTC$  can be compressed and the reward function can be effortlessly scaled. For example, if  $TTC$  is measured in seconds, taking the logarithm of  $TTC$  will transform the values to a range of -infinity to infinity, which can then be scaled and bounded to a range of 0 to 1 for use as a reward function. Second, using the logarithm of  $TTC$  can also make the reward function more sensitive to small changes in  $TTC$ , which is desirable in situations where safety is a critical concern. By taking the logarithm, a larger reward gradient can be created for small values of  $TTC$ , which can make the agent more cautious in these situations. Note that, in evaluating network safety, this study uses two assessment



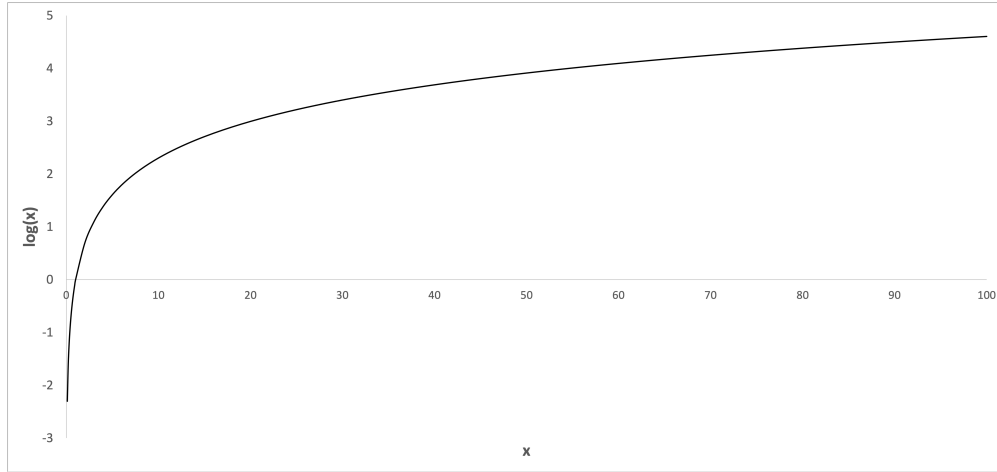


Figure 3.5: Logarithmic function

parameters to provide a wholesome safety assessment: (1) the average reward function,  $\log(TTC)$ , and (2) the number of emergency brakes.

Objective - Improving mobility:

In line with previous studies, we adopt  $TTT$  as the reward function for mobility. However, it should be noted that the focus of previous studies was to enhance the overall performance of freeway networks, and hence  $TTS$  was a more straightforward measure to design based on inflow and outflow. In contrast, defining  $TTT$  can be challenging for larger networks. We propose  $TTT$  for a network in the following manner.

$$TTT = \sum_{a \in A} \text{Current}_{TTT_a} \times N_a \quad (3.29)$$

where:

$TTT$  : total travel time for a step

$\text{Current}_{TTT_a}$  : current travel time on an edge  $a$

$N_a$  : number of vehicles on the edge  $a$

Once the agent takes an action, the simulator computes the reward function  $TTT$  for each step. A higher  $TTT$  value indicates poorer mobility because it implies that vehicles take longer to traverse the network. Therefore, in this study, the reward function based on  $TTT$  is negative.

Objective - Dampening traffic oscillation:

In section 3.3, we demonstrated how vehicles' rerouting decisions can lead to traffic oscillation in the routing links. One of the goals of this study is to mitigate traffic oscillation and investigate its impacts on the network (details in section 3.6.3). Previous studies have employed RL-based methods to mitigate traffic oscillation by using time headway as the reward function (Desjardins and Chaib-Draa, 2011; Qu et al., 2020; Zhou and Qu, 2016). However, these studies are restricted to improving traffic oscillation through vehicle-to-vehicle communication in a specific area of the network, whereas this study focuses on mitigating traffic oscillations in the routing links of the entire network. We propose a reward function for dampening traffic oscillations in the routing links in a network, which involves calculating the difference in the number of vehicles in different routes ( $\Delta\text{NumVeh}$ ) as shown below.

$$\Delta\text{NumVeh} = (\text{NumVeh}_{\text{route1}} - \text{NumVeh}_{\text{route2}}) \quad (3.30)$$

where:

$\Delta\text{NumVeh}$  : relative difference in the number of vehicles in different routes

$\text{NumVeh}$  : number of vehicles in a route

An alternative to using the number of vehicles in the routing link as a measure for improving traffic oscillation is to use density. However, density can be problematic as it is highly sensitive to the dimensions of the link, which can lead to undesired results.

### 3.5.5.3 Action:

VSLs control traffic flow by regulating speed limits, which are considered actions by TD3-based VSLCs. In real-world scenarios, speed limits are typically discrete and displayed on variable message signs with 5 mph intervals. To accommodate this, the VSLC agent is given a set of actions to choose from, including speeds of 30, 35, 40, 45, 50, 55, 60, and 65 mph. For the remaining roads without VSLC, the speed is set to a constant value of 40 mph. It's important to note that the TD3-based VSLC algorithm produces continuous actions as a probability output, not discrete actions. To convert these continuous actions into discrete speed limits with 5 mph intervals, the equation  $V_0 + I * a_t$  is used, where  $V_0$  is the initial speed of 30 mph,  $I$  is the integer constant referring to the 5 mph interval, and  $a_t$  is the action taken at time  $t$ , including added noise (see Section ?? for details).

### 3.5.6 TD3 configuration and algorithm summary

This section will summarize the TD3 algorithm as well as the parameters used. For TD3 implementation we used TensorFlow 2.12.0-rc2 and Keras 2.12.0 as deep learning API to develop the TD3 model. Different parameters have been used for the TD3 model which has been summarized in Table 3.1.

Finally, we summarize the proposed TD3 framework in Algorithm 1.

Table 3.1: TD3 Parameters

Parameters	Value
Soft replacement, $\tau$	0.005
Learning rate of critic network	0.005
Learning rate of actor network	0.002
Discount factor, $\gamma$	0.9
Exploration-exploitation parameter, $\epsilon$	0.9
Memory capacity	1000000
Batch size	256
Noise scalar, $b_t$	1.5
Noise decay	0.995
Policy learning frequency	4

### 3.5.7 SUMO configuration for the TD3 model

We used SUMO as the simulation software for our experiment. We selected SUMO as the simulation environment as it is open-source software with high flexibility, which has also been used in RL-based studies for traffic control (Boukerche et al., 2021; Kušić et al., 2021; Wang et al., 2022; Zeynivand et al., 2022). SUMO has Traffic Control Interface (TraCI) package that is useful to obtain all the necessary traffic-related parameters. For the experiment, we selected a single vehicle type: passenger car, with a length of 5 meters. Each vehicle in a fleet is assigned an individual "speedFactor", which is a speed multiplier applied to the road speed limit to achieve different desired driving speeds in a real-world situation. The "speedFactor" is set as  $norm(1, 0.1, .2, 2)$  where mean is 1, the standard deviation is 0.1, and [0.2, 2] is the default range for cut-off. This means that the speedFactor parameter will be randomly chosen from a range of values between 0.2 and 2, with a higher likelihood of being closer to 1 due to the mean value. We also choose vehicle attributes such as the car-following model and lane-change model for our vehicle type. As the car-following model, we selected the Intelligent driver model (IDM) by Martin Treiber (Treiber and Kesting, 2017) with a default setting that provides a conservative lane change gap acceptance (By default declaration rate  $> 4.5m/s^2$ ). This parameter,  $Tau$ , which addresses drivers desired minimum time

---

**Algorithm 1** Twin-delay deep deterministic policy gradient (TD3)

---

- 1: Set state, action space, and reward function in SUMO
  - 2: Algorithm parameters:  $\epsilon$ ,  $\alpha$ ,  $\beta$ ,  $\tau$ , noise parameter
  - 3: Randomly initialize two critic networks  $Q_1$ ,  $Q_2$  and actor network  $\mu$  with parameters  $\theta^{Q_1}$ ,  $\theta^{Q_2}$ , and  $\theta^\mu$  respectively
  - 4: Initialize target network weights  $\theta^{Q'_1} \leftarrow \theta^{Q_1}$ ,  $\theta^{Q'_2} \leftarrow \theta^{Q_2}$ , and  $\theta^{\mu'} \leftarrow \theta^\mu$
  - 5: Initialize replay memory  $N$
  - 6: **for** each episode **do**
  - 7:     Start traffic simulation with SUMO
  - 8:     Initialize random process using Laplacian process (Equation 3.25) for noise exploration
  - 9:     Obtain initial state  $s_1$  from SUMO
  - 10:    **for** each  $t$  of episode **do**
  - 11:       Select an action from the current policy:  $\tilde{a}_t = \mu(s_t; \theta^\mu)$  and perform the exploration method using Laplacian process (Equation 3.25) to obtain  $a_t$
  - 12:       Determine  $V_0 + I * a_t$  and feed into SUMO
  - 13:       Obtain reward  $r(s_t, a_t)$  and next state  $s_{t+1}$
  - 14:       Store a tuple  $(s_t, a_t, r(s_t, a_t), s_{t+1})$  in replay memory
  - 15:       Sample minibatch of  $N$  transitions  $(s_i, a_i, r_i, s_{i+1})$  from replay memory
  - 16:        $y_i = r_i + \gamma \min_{j=1,2} Q'_j(s_{i+1}, (\mu'(s_{i+1}; \theta^{\mu'}) + \epsilon); \theta_j^{Q'})$  where  $\epsilon \sim \text{clip}(\mathcal{N}(0, \sigma), -c, c)$
  - 17:       Update critics  $\theta^{Q_i} \leftarrow \frac{1}{N} \sum_i (y_i - Q(s_i, a_i; \theta^{Q_i}))^2$
  - 18:       **if**  $t \bmod d = 0$  **then**
  - 19:          Update  $\theta^\mu$  by the DPG
  - 20:           $\nabla_{\theta^\mu} J = \frac{1}{N} \sum_i [\nabla_a Q_1(s, a; \theta_1^Q)]_{s=s_i, a=\mu(s_i)} \nabla_{\theta^\mu} \mu(s; \theta^\mu)_{s=s_i}$
  - 21:          Update target network:
  - 22:           $\theta^{\mu'} \leftarrow \tau \theta^\mu + (1 - \tau) \theta^{\mu'}$
  - 23:           $\theta^{Q'_j} \leftarrow \tau \theta^{Q_j} + (1 - \tau) \theta^{Q'_j}$  for all  $j = 1, 2$
  - 24:       **end if**
  - 25:    **end for**
  - 26: **end for**
- 

headway, is 1 second in the model.  $\tau$  less than the reaction time leads to unsafe driving. For the lane-changing model, we used the default lane-changing model in SUMO by Jakob Erdmann (Erdmann, 2014). In SUMO "lcspeedGain" implies the eagerness among drivers to change lanes to gain speed. For our model, we choose the default value which is 1 for "lcspeedGain". For the training period, the setting of SUMO has been kept simple to observe the RL-model learning capability. We conduct sensitivity analysis on some of these parameters and observe the learning capability of the model in the later section (Section 3.6.4). One of the most important settings of

SUMO for this study is to let the vehicles re-route. The default setting of SUMO lets the vehicle choose the shortest path. In order to change that and comply with our study setting we used *traci.vehicle.rerouteTraveltime* for each vehicle which lets the vehicles choose their routing road depending on the speed of the road and takes vehicle ID (*vehID*) and current travel time (*currentTravelTimes=True/False*) as arguments. When *currentTravelTimes* is False, empty-network travel time on each edge or travel time loaded via weight files is used. On the other hand, when *currentTravelTimes* is True, the current average travel time on each edge is set as travel time. For our study, we let vehicles reroute based on the current travel time (*currentTravelTimes = True*). When vehicles are let to reroute, the average travel times in the network for all edges are collected. Once a vehicle reroutes, the edge weights (travel time of all edges) are updated. Hence when vehicles are allowed to re-route, it slows down the simulation process and a bottleneck forms at the upstream edge. (It would be interesting to see if the algorithm when safety is considered, overcomes the bottleneck) Simulation for each episode is 600 seconds with 12000 veh/hr. The number of vehicles taken in such a way to have enough vehicles for training and avoid overfitting and underfitting.

### 3.6 Experiment Results

This section presents the results of the proposed model of this study. We start by defining a base case for the study. As vehicles' routing decision-making process generates several traffic events such as more lane change, traffic oscillation, and bottlenecks, controlling the speed limit of one of the routing links (either Link 2-3 or Link 2-4), may decrease the traffic events to some extent. Therefore, as the base case, VSLC is implemented at Link 2-3 to improve mobility and safety. The state chosen for mobility and safety is the occupancy of all links. The reward for mobility and safety is *TTT* and  $\log(TTC)$  respectively as discussed in Section 3.5.5. The performance of the models is calculated from the average reward of the models across 20 episodes.

In the remainder of the section, we first analyze the network mobility and safety of the base case. Besides the base case results, the proposed model has been compared with different DRL models. In addition, the VSLC implementation for different traffic control strategies has been investigated. Finally, we conduct sensitivity analyses on the placement of VSLC in the network to improve traffic safety and mobility.

### 3.6.1 Base Case Results

In this section, we discuss the base case of our study, which will be used in the later section to investigate whether controlling one of the routing links can improve network safety and mobility. For the base case investigation, only the average safety reward ( $\log(TTC)$ ) is used to compare network safety; a more detailed discussion is provided in Section 3.6.3 and Section 3.6.5. For the base case, VSLC is implemented in Link 2-3, which is one of the routing links, that would be used to achieve two objectives: (1) improving network safety, and (2) improving network mobility. It is important to note that the study for network mobility and safety is done separately and not simultaneously. This study also observes how network mobility is affected when the VSLC is used to improve network safety, and conversely, how network safety is impacted when VSLC is used to enhance network mobility. The plot for improving network mobility and safety for both base case objectives is shown in Figure 3.6.

Figure 3.6a depicts the improvement in network safety achieved through VSLC implementation in Link 2-3 for both objectives. The results demonstrate that the TD3 model is capable of effectively enhancing network safety when the objective of the agent is to minimize traffic crash risk, as evident from the increasing trend in safety metrics post-training. However, when the objective of the agent is to improve network mobility, network safety declines progressively over time.

Figure 3.6b demonstrates the improvement in network mobility for the base case when the TD3

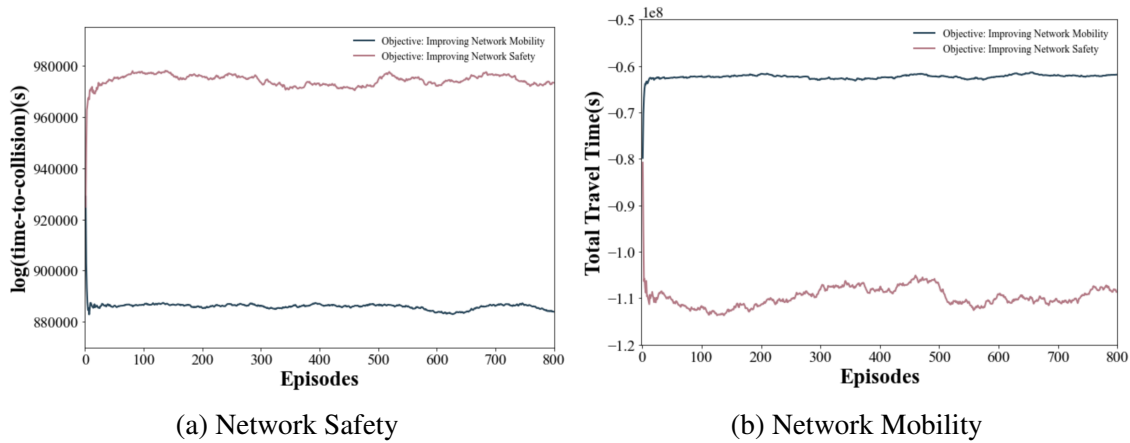
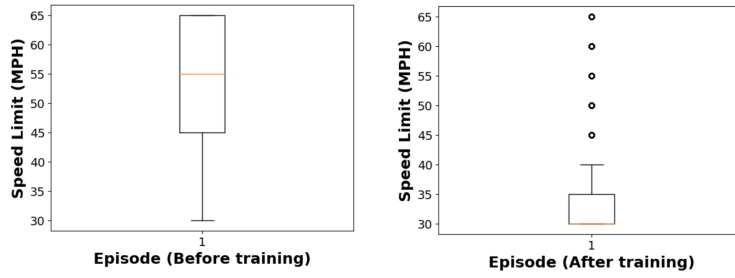


Figure 3.6: Base Case Results

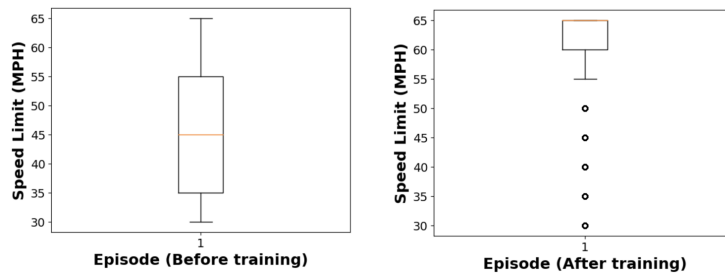
model has different objectives to achieve. The model is successful in enhancing network throughput when the reward is set to reduce total travel time in the network. Conversely, when the objective is to improve safety, the network mobility deteriorates over time. Notice that, the model set to improve network safety exhibits several oscillations, while the improvement in network mobility reaches stability quickly. The reason behind this is the assumption of full observability in the MDP process, where the agent can observe the environment states fully and the reward under a given state, which can be controlled by the agent's action. The agent can observe the occupancy of each link and provide a speed limit that facilitates fast vehicle passage for network mobility to improve. However, for the safety improvement case, the interaction between the vehicles plays a significant role that is beyond the agent's control, resulting in oscillation in the learning process to enhance network safety over time.

The action of the agent provides further insight into the base case network safety and mobility. In Figure 3.7, the actions undertaken by the agent for enhancing safety and mobility are illustrated before and after undergoing training. The figure provides a comprehensive portrayal of the variations in the agent's behavior following the training process, thereby enabling several informative





(a) Network Safety



(b) Network Mobility

Figure 3.7: Speed Limit for the Base Case

insights. In Figure 3.7a, it can be observed that prior to the training process, the range of speed limits adopted by the agent to enhance network safety extended from 30 to 65 mph, with a higher frequency of actions taken in the 45 to 65 mph range. Despite being advantageous for network mobility, this higher speed limit resulted in suboptimal safety performance, as evidenced by the initial episodes of Figure 3.6a, and higher mobility performance, as observed in Figure 3.6b. Subsequent to the training process, the agent exhibited a superior level of network safety improvement, with the speed limit ranging from 30 to 40 mph, which resulted in enhanced safety performance (Figure 3.6a) but reduced mobility performance (Figure 3.6b).

Similarly, prior to the training process, the range of speed limits employed by the agent to improve network mobility extended from 30 to 65 mph, with a higher frequency of actions taken in the 35 to 55 mph range. This lower speed limit was advantageous for improving network safety, as

evidenced by the higher initial safety performance in Figure 3.6a, but resulted in lower mobility performance, as seen in Figure 3.6b. Post-training, the agent exhibited superior network mobility improvement, with the speed limit ranging from 55 to 65 mph, resulting in decreased safety performance (Figure 3.6a) but improved mobility performance (Figure 3.6b).

### 3.6.2 Comparison with other Methods

In this section, we take the base case scenario and compare the proposed algorithm with different algorithms, both value-based (i.e., DQL) and policy-based algorithms (i.e., soft actor-critic model (SAC), DDPG). To understand the potential of each algorithm to improve network mobility and safety, we compare them against no VSL control. Here, network safety is measured using the  $\log(TTC)$  and network mobility using total travel time as discussed in detail in Section 3.5.5. All models are trained for 800 episodes with the same SUMO parameters. Before diving into the comparisons, we first discuss all the models.

- **NoVSLC:** No VSLC is deployed in the network. The speed limit for each edge is 40 mph.
- **VSLC-DQL:** DQL performs best for discrete action space where the action is taken by finding the action maximizes Q-value. Since we have continuous action space, we define a neural network that will take state and output action. The structure of the neural network is similar to that of the actor network of the proposed model. The parameters are kept the same as the proposed model to guarantee the same learning capability.
- **VSLC-SAC:** The SAC model has two critic networks and an actor network as the proposed model does. In addition, it has a value network.
- **VSLC-DDPG:** The DDPG is model is similar to the proposed TD3 model with same parameters. The difference between the two models is DDPG has only one critic network and the

weight update mechanism is different from the proposed model.

- **VSLC-TD3:** This is the proposed model. See Section 3.5 for details.

Table 3.2: Average Performance of Different Models in Improving Network Safety and Mobility in Base Case

Method	Network safety (unit: seconds)	Network mobility (unit: seconds)
No VSLC	9.5262e+05	9.1878e+07
VSLC-DQL	8.8534e+05	1.1302e+08
VSLC-SAC	9.2121e+05	8.4903e+07
VSLC-DDPG	<b>9.9546e+05</b>	7.2324e+07
VSLC-TD3	9.8919e+05	<b>6.1384e+07</b>

Table 3.2 shows the average performance of different methods across the 100 test episodes for the base case. The result shows that by implementing VSLC in Link 2-3 using different algorithms, most methods can improve network safety and network mobility. For improving network safety, VLSC implemented using DDPG can provide the best, which improves traffic safety 4.50% higher than no VSLC control. We proposed TD3-based VSLC that addresses the overestimation issue faced by DDPG, which improves network safety by 3.84 % more than no VSL control. The proposed algorithm also provides the best case for network mobility which is 33.19% higher than no VSL control. Using DDPG network mobility can also be improved significantly which is 21.28% more than no VSL control. VSLC implementation using SAC can improve network mobility by 7.59% higher than no VSL control, however, network safety cannot be improved using this method (3.29 % lower than no VSL control). TD3 and DDPG seem to have provided a better result in improving network mobility and safety than SAC. This could be because the deterministic policies (DDPG and TD3) can be more stable and easier to optimize than the stochastic policy used in SAC. The deterministic policies allow for a straightforward exploration of the action space because the agent can learn a specific action for each state. On the other hand, the stochastic policy of SAC

may require more exploration to learn a good policy because the agent must sample from the policy to select an action. Contrary to the policy-based methods, the value-based method (DQN) fails to improve network safety and mobility, which is 7.05% and 23% lower than no VSL control respectively. This could be because value-based algorithms in continuous action spaces require discretization of the action space, which can lead to a large number of possible actions, making the learning process very slow and computationally expensive. Additionally, discretization can lead to suboptimal actions being chosen, as the algorithm is limited to a finite set of actions.

### 3.6.3 Assessment of Different Traffic Control Strategies in the Network

We compare the routing of vehicles using four different traffic control strategies: no VSLC, the VSLC implementation for mobility and for safety, and traffic engineering practice. The traffic control using the VSLC implies that the speed limit of one of the routing links (Link 2-3 in Figure 3.1) is controlled using VSL while vehicles are let to choose their route. Traffic engineering practice implies that traffic engineers would control the flow in the routing link in such a manner that both routes are equally used. Here, traffic engineering practice is implemented in this study by changing the speed limit of Link 2-3 in such a manner that by fluctuating the speed limit of Link 2-3 (30mph to 65mph) and keeping the speed limit of Link 2-4 constant (40 mph), the traffic engineers would ensure the flow in both the routing links are continuous and fewer traffic oscillations in the routing links. It should be noted that, in all scenarios considered, vehicles have the freedom to choose their routes based on the current travel time to reach their destination. To investigate the impact of different control strategies in the network, we compare the average performance of different traffic control strategies in terms of routing, traffic mobility, and safety measures in Table 3.3. Note that we choose the best records for each category to conduct the comparison as determining the average of some categories (e.g. routing) would not reflect the actual scenario for the control strategies. In our experiment, VSLC implementation for mobility achieves its best mobility measure in the

796th episode. VSLC implementation for safety and traffic engineering practice achieved their corresponding best performance in the 798th and 795th episode respectively. We provide more insights into the comparison of different categories below.

Table 3.3: Average Performance of Different Traffic Control Strategies

	Routing		Lane change frequency (%)	$\sum \log(TTC)$	$\log(TTC)$ due to Lane change	Total travel time (sec)
	% of vehicles in Route 1	% of vehicles in Route 3				
No VSL	49.41	50.59	13.97	9.526e+05	6.27E+04	-9.188e+07
VSLC implementation (mobility)	64.35	35.65	10.13	9.149e+05	5.36E+04	-6.576e+07
VSLC implementation (safety)	26.38	73.62	14.25	1.002e+06	6.32E+04	-1.349e+08
Traffic Engineering Practice	50.32	49.68	26.66	1.005e+06	9.81E+04	-1.534e+08

## Routing

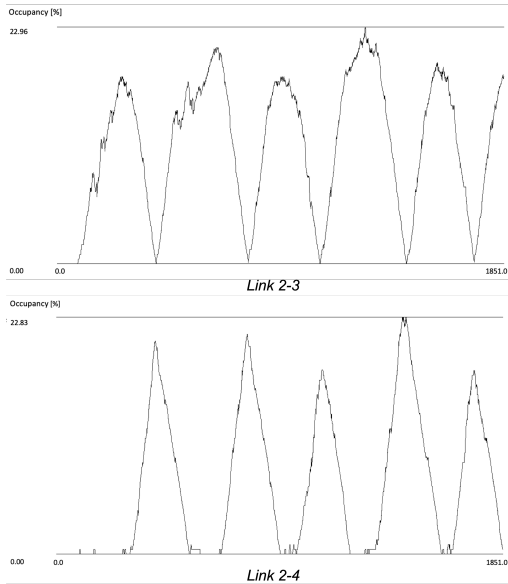
In this study, we evaluate and compare the effectiveness of various traffic control strategies on vehicle routing in the network. It can be seen that by implementing traffic engineering practices using the Equation (3.30), nearly equal occupancy in both routes can be achieved (Table 3.3). This can be further seen in Figure 3.8 where the occupancy of Link 2-3 and Link 2-4 is shown for a period of time for different traffic control strategies. Figure 3.8c illustrates that vehicles utilize both routes simultaneously, as indicated by the continuous occupancy of both routes throughout the entire simulation period. The speed profile for traffic oscillation shows how traffic engineering practice accommodates equal routing by providing speed in the Link 2-3 that changes widely over the simulation period time (Figure 3.10). Note that, this strategy helps dampen traffic oscillation in the routing links. Without VSLC, vehicles tend to use both routes in a nearly equal manner. However, Figure 3.2 indicates that as the number of vehicles on one route reaches a peak (maximum occupancy at that time), the number of vehicles on the other route drops to a valley (minimum oc-

cupancy at that time), with no vehicle taking the latter route during that time period. This suggests that vehicles alternate between the routes, resulting in traffic oscillation in the routing links, which is more pronounced with the duration of occupancy being equal in Figure 3.2.

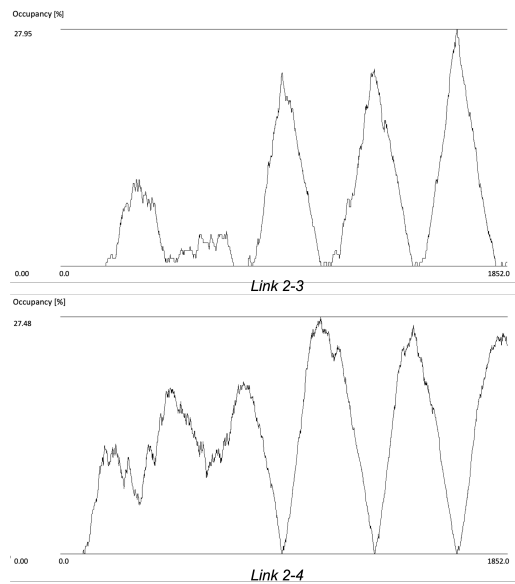
By implementing VSLC to improve mobility, it can be seen that more vehicles use Route 1 where the VSLC has been placed. This can be seen in Figure 3.8a where the duration of occupancy of Link 2-3 is more than Link 2-4. This is because the VSLC provides a higher speed limit (Figure 3.10) in Link 2-3 which attracts more vehicles on the route. This strategy demonstrates a clear oscillation (wave) that takes place on the two edges. VSLC implementation for safety shows more vehicles using route 2 than Route 1 which can also be seen in Figure 3.8b. The speed profile for this strategy shows (Figure 3.10) a lower speed limit is provided in Link 2-3 in order to improve safety which discourages the vehicles to use Link 2-3 and choosing Link 2-4 more for routing.

### **Network Mobility**

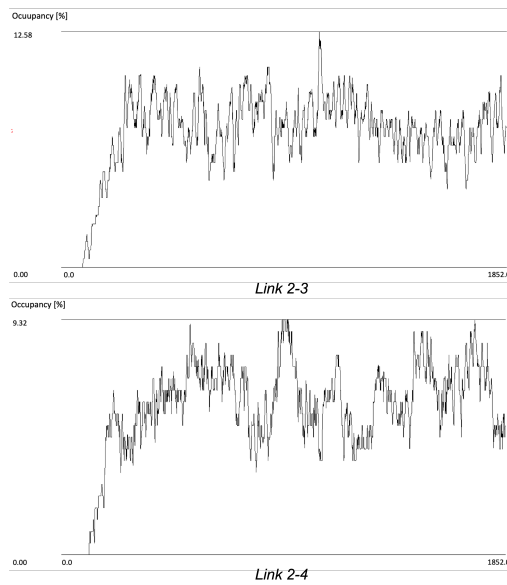
Table 3.3 shows the traffic mobility measure (*TTT*) of all the traffic control strategies. It shows that without VSLC traffic would have a longer time to traverse through the network. VSLC implementation for mobility improvement provides the highest mobility where it improves traffic throughput by 28.4 % higher than without any VSLC. This is because VSLC in order to improve mobility provides a higher speed limit (Figure 3.10) which helps vehicles with the decision-making process of choosing routes and forms less bottleneck in Link 1-2 as shown in Figure 3.9 in average queue length. Contrary to that, VSLC implementation for improving safety worsens traffic mobility which is 46.82% lower than no VSL control. This is because this strategy provides a lower speed limit (Figure 3.10) and the speed difference between the routes is lower. This makes the routing decision-making process harder for the vehicles forming longer bottlenecks as shown in Figure 3.9a. The traffic engineering practice strategy is the worst strategy for improving mobility as the goal of this strategy is to balance the occupancy in both routes. As VSLC provide a varied range of speed limit over the simulation period to balance vehicles in the routes, vehicles need



(a) VSLC implementation (mobility)



(b) VSLC implementation (safety)



(c) Traffic Engineering Practice

Figure 3.8: Traffic oscillation when controlled

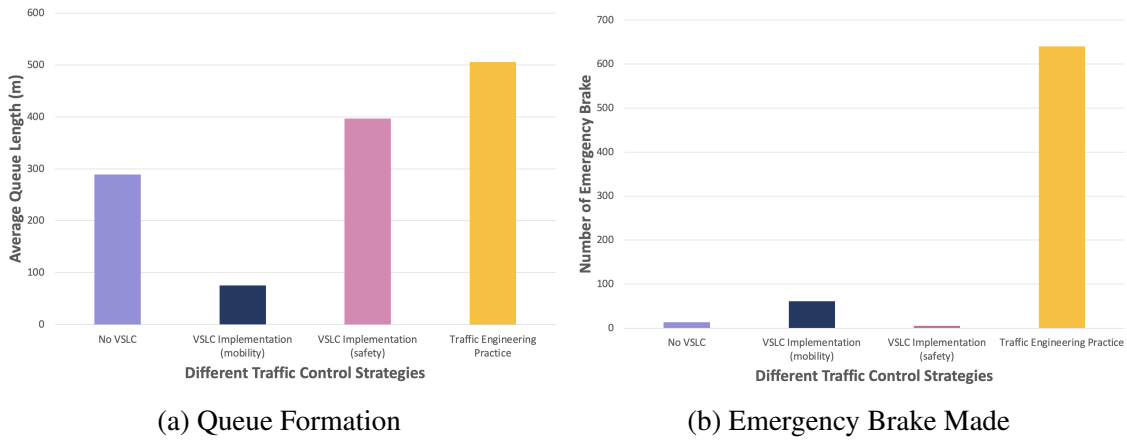


Figure 3.9: Traffic oscillation when controlled

to constantly decide to choose routes forming a large queue in Link 1-2 (See Figure 3.9a). This provides the worst-case for network mobility improvement.

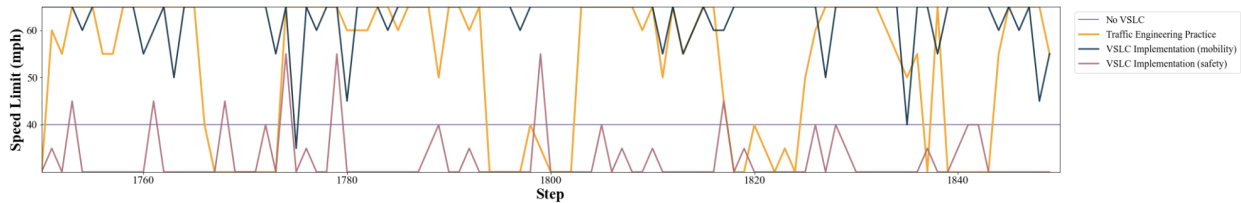


Figure 3.10: Speed profile of Link 2-3 for different traffic control strategies

## Network Safety

Table 3.3 shows the traffic safety measure ( $\log(TTC)$ ) of all the traffic control strategies. It shows that network safety measure is slightly higher for traffic engineering practice than when VSLC is implemented to improve network safety. One of the reasons behind that is the number of lane change frequency is higher than the other strategies (26% of all vehicles), where it contributes to



9.77% of the total  $\log(TTC)$ . However, for traffic engineering practice, the number of emergency brakes pressed during the simulation is the highest (Figure 3.9b). This means that this traffic control strategy causes the most emergency situation for the vehicle traversing the network. This is because the fluctuating speed limits imposed by Traffic Engineering standards, ranging from 30 to 65 mph, can cause significant disruptions in vehicle speed, leading to abrupt braking situations. Contrary to that, VSLC implementation for safety improvement not only provides a high safety measure but also has the least number of emergency situation risen during simulation. The speed limit provided by the VSLC for this strategy also contributes to promoting safety as the overall speed limit for the whole network is lower than other traffic control strategies. VSLC implementation for mobility provides a poor improvement in  $\log(TTC)$  which is 3.95% lower than no VSL control. Overall, VSLC implementation for safety improvement is the best-case scenario for safety improvement among all the traffic control strategies.

#### 3.6.4 Assessment of Generalization of the Agent

Assessing the generalization of an RL agent is essential because it determines the agent's capability to apply its learned behavior to novel and different situations that were not encountered during its training. A generalizable agent can adapt to new and unseen scenarios and demonstrate effective performance even when environmental conditions have changed. In the context of road traffic, evaluating the transferability of an RL agent is particularly critical due to the complex, dynamic, and ever-changing nature of traffic environments. The variation in traffic conditions and unforeseen events in real-world traffic scenarios can cause the failure of VSL control (Wang et al., 2012). In this study, the agent is trained in a simulated environment (See details in Section 3.5.7). However, it is important to evaluate whether the trained model is capable of performing well in real-world traffic situations with varying conditions. While it may not be feasible to fully test the model in a real-world deployment, we can still evaluate its transferability by examining its performance under

various traffic scenarios with different attributes that are distinct from the ones it was trained on.

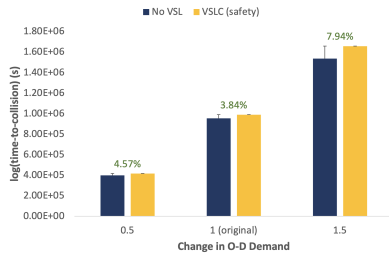
As mentioned in Section 3.5.7, the model is trained on a fixed demand of 12000 veh/hr for 600 seconds with a "speedFactor" and "lcspeedGain" value of norm(1,0.1,0.2,2) and 1, respectively. To test the generalization of the algorithm, we evaluate the agents in different environments with varying demands, mean and deviation values of "speedFactor", and "lcspeedGain". While assessing the model for one of these attributes, the rest of the attributes are kept constant. We evaluate the performance of all models on 100 test episodes. Note that to assess network safety improvement for different traffic parameters, we are using the average reward function defined for network safety.

#### **Origin-Destination Demand:**

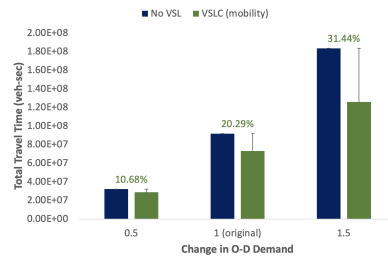
Figure 3.11a shows that as the O-D demand increases the safety of the network increases when the same speed limit is imposed for all links (no VSL control) and with VSLC imposed speed limit. When the O-D demand is less (0.5 times of original O-D demand), both with and without VSLC have nearly the same network safety measure. However, the difference between the safety measures with and without VSLC is more when the O-D demand is 1.5 times the actual O-D demand used (7.94%). This implies that using VSLC to improve network safety is less significant when the O-D demand is less. The use of VSLC has more value when the O-D demand is high. Contrary to network safety, network mobility decreases as the O-D demand increases (Figure 3.11b). Without VSLC, network throughput decreases significantly with the increase in O-D demand (total travel time increases from  $3.22 \times 10^7$  to  $1.84 \times 10^8$  veh-sec) which can be improved using VSLC ( $2.878 \times 10^7$  to  $1.26 \times 10^8$  veh-sec)

#### **Mean value of speedFactor:**

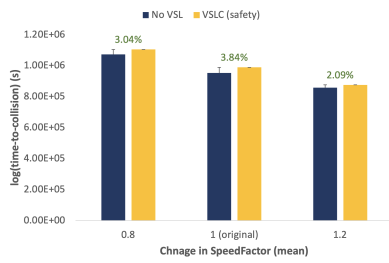
In SUMO, "speedFactor" is used to estimate the desired free-flow driving speed of the vehicle by multiplying the factor by the road speed limit. If the mean of "speedFactor" is greater than 1,



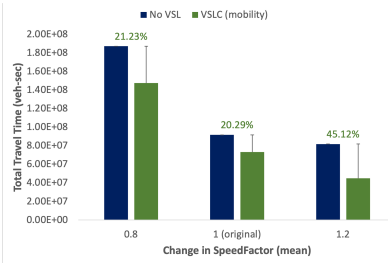
(a) The overall network safety measure: O-D demand



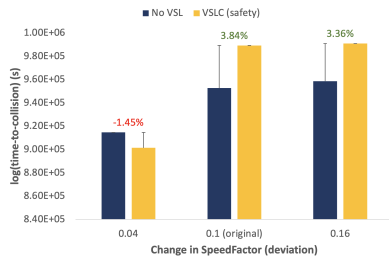
(b) The average travel time: O-D demand



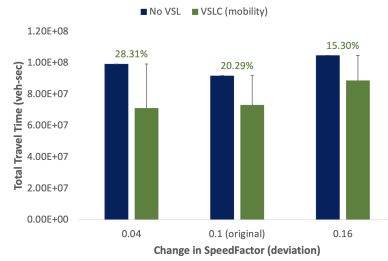
(c) The overall network safety measure: speedFactor (mean)



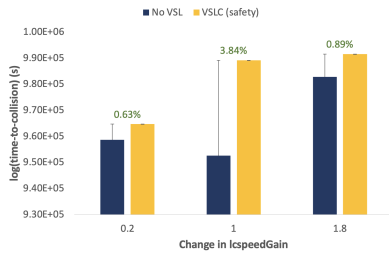
(d) The average travel time: speedFactor (mean)



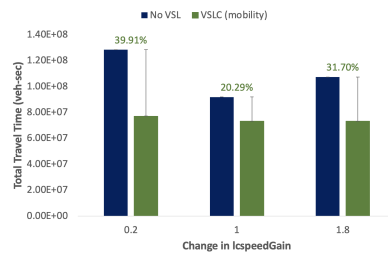
(e) The overall network safety measure: speedFactor (deviation)



(f) The average travel time: speedFactor (deviation)



(g) The overall network safety measure: lcspeedGain



(h) The average travel time: lcspeedGain

Figure 3.11: Effect of Different Traffic Attributes on Network Safety and Mobility

vehicles can exceed the edge speed limit. Figure 3.11c and 3.11d show the impact of different values for the mean of "speedFactor" on network safety and mobility respectively. It is evident that as the mean value of "speedFactor" increase, the network safety measure decreases (Figure 3.11c), and the network mobility measure increases (Figure 3.11d). In a real-world situation, the desired driving speed can differ among vehicles in a fleet. If vehicles can achieve their desired speed by exceeding the speed limit posted on the road, this will help vehicles traverse the network quickly. However, this will have an adverse impact on network safety as speeding can result in traffic incidents. Figure 3.11c shows that although network safety decreases with increase in the mean of "speedFactor", using VSLC can provide higher network safety than no VSL control. Similarly, VSLC implementation can improve network mobility significantly in comparison to no VSL when vehicles exceed the speed limit (Figure 3.11d).

#### **Deviation of speedFactor:**

As previously stated, the "speedFactor" parameter is used to adjust the free flow speed of vehicles on a link by multiplying it with the speed limit. Meanwhile, speed deviation refers to the variation in desired driving speed among the vehicles on a link. Figure 3.11e depicts that with the increase of speed deviation among vehicles, network safety improves for both with VSLC and no VSL control. A small amount of speed variation in the traffic flow would indicate that the vehicles are driving close to the same speed limit. This implies that the vehicles do not need to frequently adjust their speed to match the speed of other vehicles, leading to a smoother traffic flow and enhanced safety. However, for a speed deviation value of 0.04, it can be seen that no VSL control improves network safety better than using VSL control. This means that since vehicles are closely following the recommended speed, using a constant speed limit is favorable; implementing variable speed limit in this case might increase traffic crash likelihood. On the other hand, an increase in the deviation of desired driving speed decreases traffic mobility. This is because, when drivers are traveling at different speeds, it can create bottlenecks and congestion on the roadway. Faster drivers may

be forced to slow down or brake abruptly, while slower drivers may be unable to keep up with the flow of traffic. This can lead to "stop-and-go" traffic patterns that decrease overall mobility. Additionally, drivers who are traveling at different speeds may need to change lanes in order to pass slower vehicles or avoid faster ones. This can lead to weaving and lane changing maneuvers that can slow down traffic.

### **lcSpeedGain:**

As stated earlier, "lcSpeedGain" refers to drivers' eagerness to change lanes in order to gain speed. Figure 3.11g illustrates that with an increase in the number of drivers changing lanes, network safety is improving. This is because drivers can avoid potential crashes on the road by changing lanes in case of emergency situations, which helps improve network safety. On the other hand, the increase in the number of drivers changing lanes decreases network mobility (Figure 3.11h). This is because, vehicles that are changing lanes require more space and can disrupt the flow of traffic, which can decrease the number of vehicles that can fit on the roadway at any given time. In addition, frequent lane changes by drivers can cause congestion on the roadway. This is because, when drivers are constantly weaving in and out of traffic, it can create bottlenecks and slow down traffic, which can decrease overall mobility.

### 3.6.5 Sensitivity of VSLC placement in the Network

Numerous studies have demonstrated the potential of implementing VSLC to enhance both the safety and mobility performance of transportation networks, especially for local issues such as segments of freeways. While VSL control has the potential to improve both safety and mobility performance in transportation networks, implementing VSL control in all links of a network can have some disadvantages. One major disadvantage is the high cost associated with installing and maintaining VSL control infrastructure throughout the entire network. This includes the cost of

installing sensors, communication equipment, and displays, as well as ongoing maintenance costs for the equipment. Another disadvantage is the potential for driver confusion and information overload due to constantly changing speed limits. If speed limits change frequently and without clear reason, drivers may become frustrated and ignore the VSL system altogether, potentially increasing the risk of accidents. This study aims to determine the optimal position in the network for VSLC implementation, which can improve network mobility and safety while considering the associated implementation cost. To achieve this, we selected a small network for examination of traffic behavior under VSL control and explored various VSLC placement options (Figure 3.12). Figure 3.12(a) shows VSLC implementation on the Link 1-2, which is utilized by vehicles prior to the point where they need to choose a route. Figure 3.12(b) controls the speed of a routing link to improve network mobility and safety, which is the base case of the study. Figure 3.12(c) and (d) assess whether dividing the links (Link 1-2 and 2-3 respectively) into upstream and downstream and controlling their speed by VSLC can further improve the network mobility and safety. The reasoning behind this approach is that if there is congestion downstream of a link, increasing the speed limit throughout the entire link could exacerbate the queue. In such scenarios, setting a higher speed limit downstream and a lower speed limit upstream of a link may mitigate the problem. In this study, we first examine the placement of VSLC for enhancing network safety, followed by a discussion on the placement of VSLC for improving network mobility.


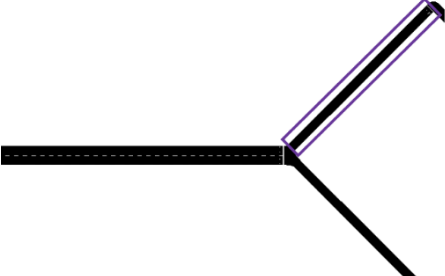
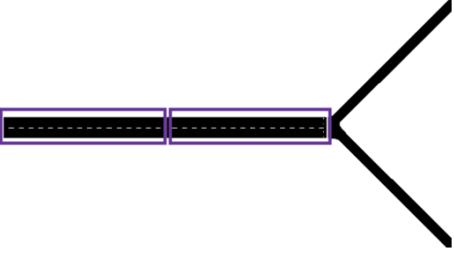
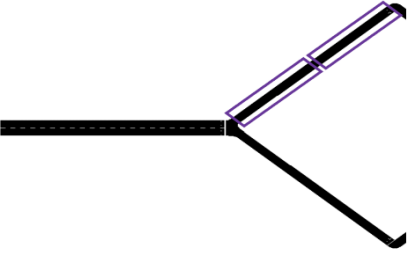
Position of VSL	Diagram
(a) One VSL in the Link 1-2	
(b) One VSL in the Link 2-3 (Base case)	
(c) Two VSLs in the upstream and the downstream of the Link 1-2	
(d) Two VSLs in the upstream and the downstream of the Link 2-3	

Figure 3.12: Position of VSL for improving network safety and mobility

### Network Safety

Figure 3.13 displays the average safety reward over a duration of 800 episodes. Notably, the base case (Figure 3.12(b)) is not the most effective in terms of improving average safety reward.

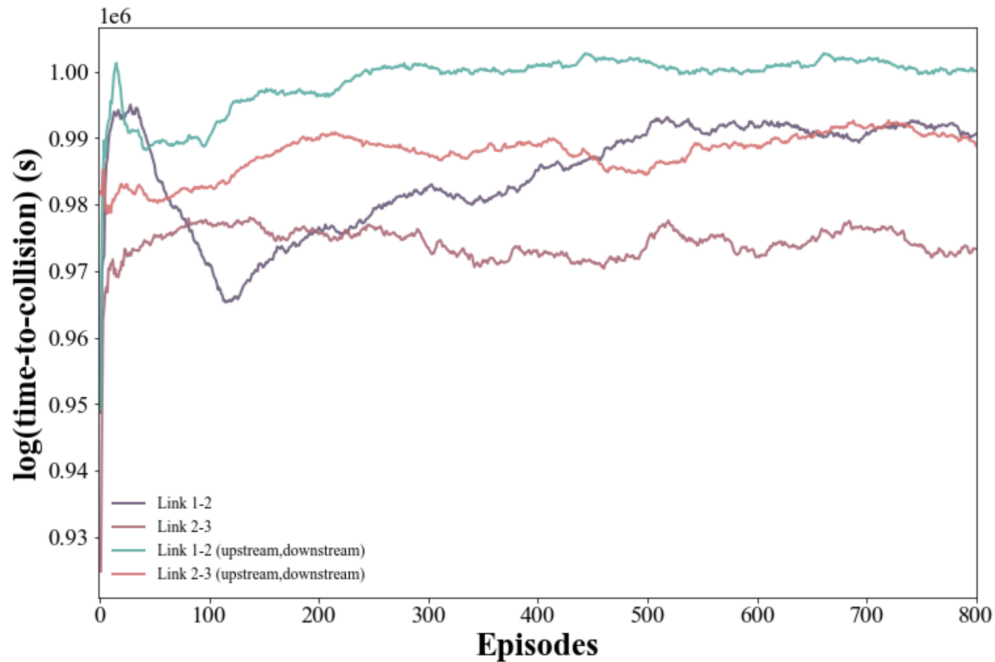


Figure 3.13: Performance of different VSLC placements to improve network safety

Figure 3.13 indicates that the VSLC placement upstream and downstream of Link 1-2 (Figure 3.12(c)) yields the most desirable outcome for improving the average safety reward ( $\log(TTC)$ ). VSLC placement in the upstream and downstream links of Link 2-3 (Figure 3.12(d)) and VSLC placement in Link 1-2 (Figure 3.12(a)) also provide a higher average safety reward, with the former reaching a steady state more quickly than the latter. In contrast, the base case for safety is the least effective in terms of improving average safety reward. However, to gain a better understanding of the differences between the various strategies and their overall safety impact on a network, a deeper investigation into the underlying traffic behavior that occurs is necessary.

Due to the vast amount of trained data, compiling traffic behavior for all episodes can be a time-consuming task. To analyze the traffic behavior of the different strategies for improving network safety, a subset of 15 episodes was randomly selected from a large amount of trained data to ensure



Table 3.4: Average Performance of Different VSLC Placement for Network Safety Improvement

	Routing		Lane change frequency (%)	Network mobility		Network safety	
	Route 1 (%)	Route 2 (%)		Total Travel Time of all vehicles (sec)	Delay due to Lane change (sec)	$\sum \log(TTC)$	$\log(TTC)$ due to Lane change
Link 1-2	50.34 (2.97)	49.66 (2.97)	16.22 (2.43)	1.220e+08 (1.651e+07)	3.193e+05 (7.9723e+04)	9.936e+05 (2.1186e+04)	7.065e+04 (5.875e+03)
Link 2-3	32.27 (7.17)	67.73 (7.17)	13.20 (2.51)	8.778e+07 (1.7735e+07)	2.626e+05 (8.4913e+04)	9.799e+05 (1.5826e+04)	6.696e+04 (5.225e+03)
Link 1-2 (upstream, downstream)	49.77 (2.77)	50.23 (2.77)	16.92 (3.31)	9.774e+07 (1.6262e+07)	3.182e+05 (1.0164e+05)	9.952e+05 (2.5877e+04)	6.709e+04 (7.341e+03)
Link 2-3 (upstream, downstream)	43.99 (5.84)	56.01 (5.84)	19.26 (6.71)	1.228e+08 (3.2419e+07)	3.585e+05 (1.6093e+05)	9.891e+05 (3.3130e+04)	8.455e+04 (1.8834e+04)

unbiased results. The average performance, traffic behavior, and speed limit of each strategy were compiled and presented in Table 3.4, Figure 3.14, and Figure 3.15, respectively. The base case, which is the least effective strategy for improving the average safety reward in Figure 3.13, was selected as a starting point for the analysis. Figure 3.15b shows that placing VSLC in Link 2-3 results in an average speed limit ranging from 30 to 35 mph during peak periods. However, due to the lower speed limit on this route compared to Route 2, the majority of vehicles (on average 67.73%) prefer to use Route 2, which has a constant speed limit of 40 mph, to reach their destination. This leads to the highest improvement in network mobility compared to all other strategies while the objective is to improve network safety, as indicated by the mobility measure of  $8.778 \times 10^7$  sec for all vehicles on average, and the lowest queue formation of 306.8 meters on average (Figure 3.14c). This is because the majority of vehicles use Route 2 and takes less amount of time to make decisions to choose a route, thereby experiencing fewer bottleneck. Additionally, there is a high number of aggressive drivers who change lanes to gain an advantage, leading to emergency situations that require the use of emergency brakes (on average 15 times) (Figure 3.14).

The placement of VSLC in the upstream and downstream of Link 1-2 (Figure 3.12(c)) yields the

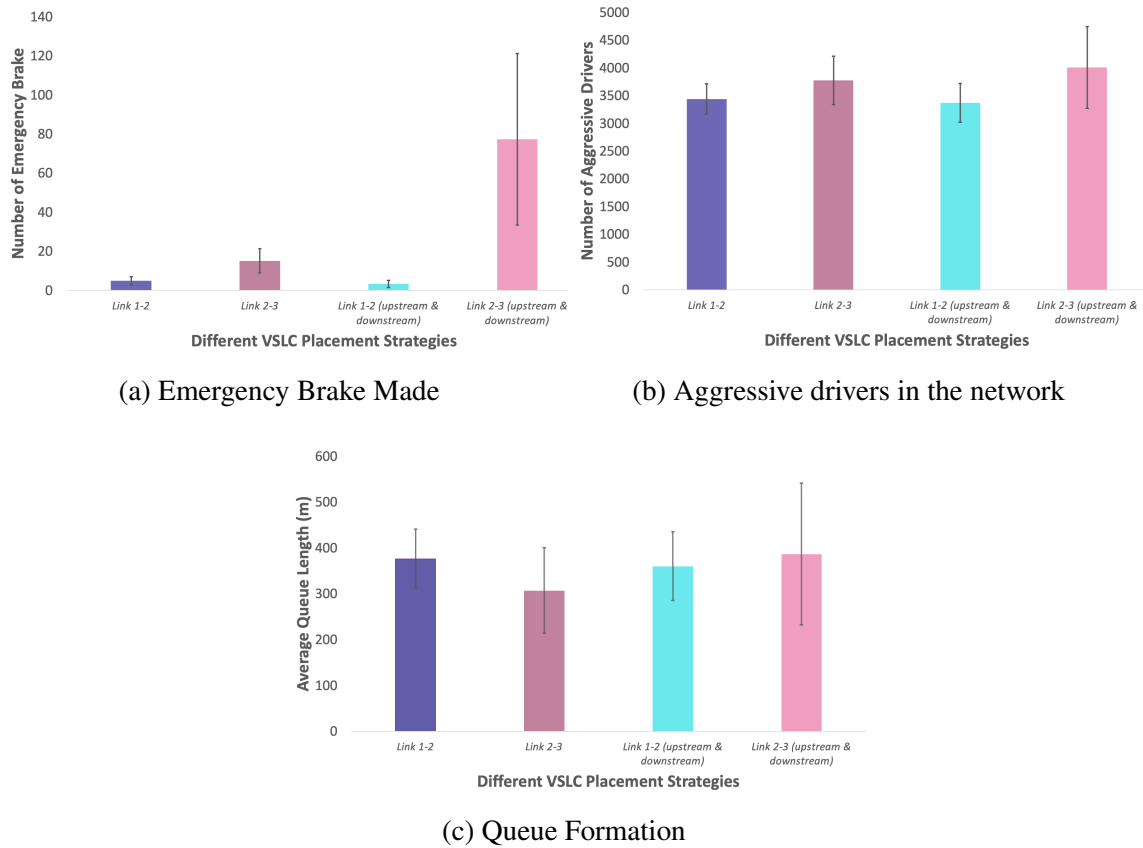


Figure 3.14: Traffic behaviors for different VSLC placements for Network Safety Improvement

highest improvement of average safety reward (1.56%) compared to the base case in the analyzed subset of episodes. The average queue formation is higher than the base case (on average 360 meters). This is attributed to the lower speed limit, ranging on average from 30 to 35 mph upstream and 35 to 40 mph downstream. These lower speed limits, however, do not pose any significant safety risks and contribute to reducing bottlenecks to some extent by limiting speeds upstream compared to downstream. The lane change frequency for this strategy is higher than the base case at 16.9% on average, accounting for 15.15% of  $\log(TTC)$  on average. Due to the reduced variance in the speed limit provided by VSLC, the number of aggressive drivers who change lanes to gain speed is 10.71% lower than the base case (Figure 3.14b), resulting in the lowest number of

emergency brakes being pressed compared to all other strategies, with an average of 3.26 times. As the routes have the same speed limit, vehicles use both routes equally.

The placement of VSLC in Link 1-2, as shown in Figure 3.12(a), results in a better improvement in average safety reward in comparison to the base case. The difference in this strategy and Figure 3.12(c) lies in the homogeneous speed limit of Link 1-2, which averages between 35 to 40 mph. This leads to a higher average queue formation compared to the previous strategy, with an average of 376 meters, which is 22.6% higher than the base case. The lane change frequency is lower compared to the previous strategy, averaging 16.9% of all vehicles, resulting in a contribution of 17.6% to total  $\log(TTC)$  on average. The reduced variance in the speed limit provided by VSLC leads to a decrease of 8.88% in the number of aggressive drivers changing lanes to gain speed compared to the base case, resulting in an average of 4.88 emergency brake activations. Similar to Figure 3.12(c), vehicles choose to use both routes equally.

VSLC placement in both the upstream and downstream of Link 2-3 (Figure 3.12(d)) also provides a higher average safety reward improvement. The speed limit varies upstream and downstream, whereas in the upstream the speed limit varies mostly between 45 to 50 mph, and the downstream has a dramatic speed limit range that varies between 30 to 40 mph. Since the upstream has a higher speed limit than the downstream, a queue forms in Link 2-3. As the speed limit on route varies widely from high to low, the decision making process to choose a route gets difficult for the drivers in Link 1-2 which causes a bottleneck. This is why the average queue formation is the highest among all the strategies (on average 386 meters) which is 26% higher than the base case. Even though Route 1 at times can provide a higher speed limit, the differential speed limit in the upstream and downstream and the queue formation make it a less favorable route to travel leading 56% of vehicles on average to travel using Route 2 more. Additionally, the lane change frequency is the highest on average among all strategies, which contributes to the average safety reward being higher (13.6% of total  $\log(TTC)$  on average). The abrupt and differential changes in speed

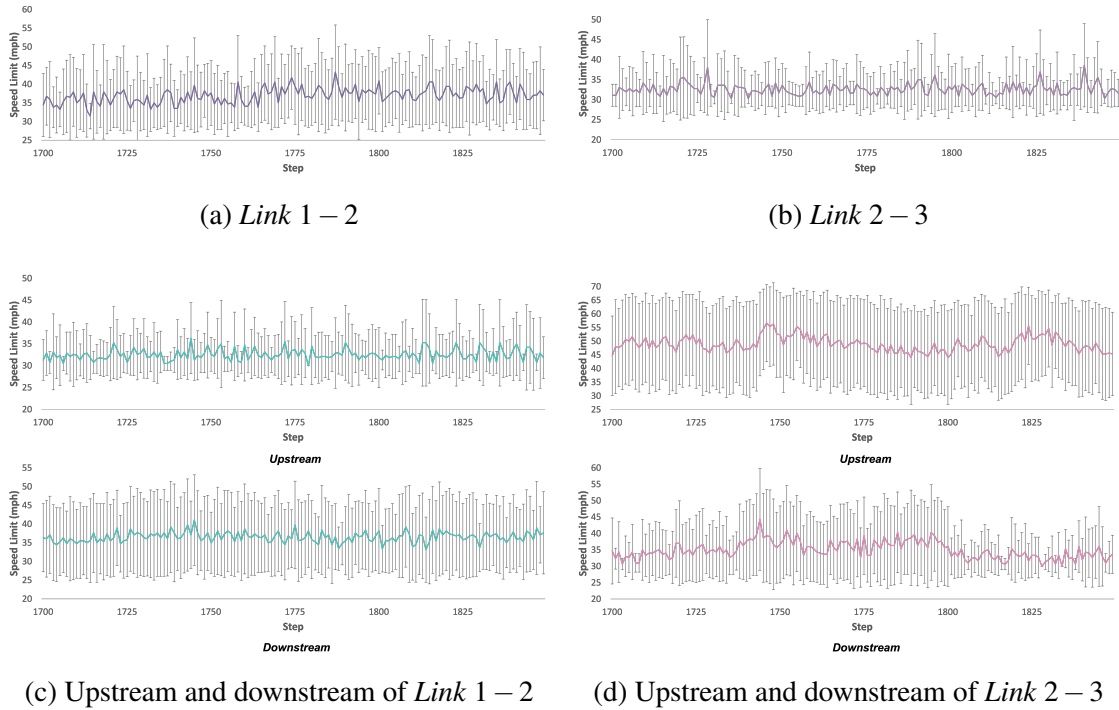


Figure 3.15: Speed profile for different VSLC placements to improve network safety

limits in the upstream and downstream areas may not be as effective, as the number of aggressive drivers tends to be the highest on average, leading to the highest number of emergency brakes being pressed compared to the base case (an average increase of 411%). Placement of VSLC in both the upstream and downstream of Link 2-3 although provides a higher safety reward during training, this strategy results in the highest number of emergency situation in the network, making the strategy less desirable for safety improvement. To summarize, in order to assess the network safety improvement, consideration of the two parameters, average safety reward ( $\log(TTC)$ ), and the number of emergency brakes are important. Using these two parameters to assess network safety, it can be seen that placing VSLC in Link 1-2 using either of the strategies (one VSLC or two VSLC upstream and downstream) provides higher network safety where drivers are not aggressively driving and fewer emergency situation arises and the average safety measure is highest.

Since the speed limit in these strategies is of a lower range, average queue formation is high which leads to poorer network mobility. On the other hand, placing VSLC in Link 2-3 using either of the strategies (one VSLC or two VSLC in upstream and downstream) leads to a higher number of aggressive drivers in the network which leads to a higher number of emergency brakes pressed. Therefore using this strategy might not always help in improving network safety.

### **Network Mobility**

The average mobility reward over 800 episodes is presented in Figure 3.16. The results indicate that the base case (Figure 3.12(b)) performs better in terms of improving network mobility than other strategies. VSLC placement in Link 1-2 (Figure 3.12 (a)) and VSLC placement in the upstream and downstream of Link 1-2 (Figure 3.12 (c)) exhibit similar levels of improvement, with the former achieving stability faster than the latter. Conversely, VSLC placement in the upstream and downstream of Link 2-3 (Figure 3.12(d)) appears to result in the worst network mobility improvement. Similar to the safety assessment, 15 episodes are randomly selected for each strategy to compare their performance, traffic behavior, and speed limit, as summarized in Table 3.5, Figure 3.17, and Figure 3.18, respectively.

The analysis of mobility improvement starts with the base case, which is identified as the most effective strategy for improving network mobility. The average speed limit achieved through the placement of VSLC in Link 2-3 is presented in Figure 3.18b, where the speed limit ranges from 60 to 65 mph over the subset of episodes. This higher speed limit in Route 1 aids drivers in making decisions more efficiently, resulting in 65% of vehicles opting for Route 1. Consequently, the average queue formation is the lowest, at an average of 140 meters. Additionally, the total delay resulting from this strategy is the lowest among all the strategies, with a value of  $7.166 \times 10^7$  seconds for all vehicles, where lane change activity contributes only 0.18% of the total delay. Moreover, this strategy results in the highest  $\log(TTC)$  among all the strategies, with lane change activity contributing 6.12% of the total  $\log(TTC)$ . Furthermore, due to the reduced variance in

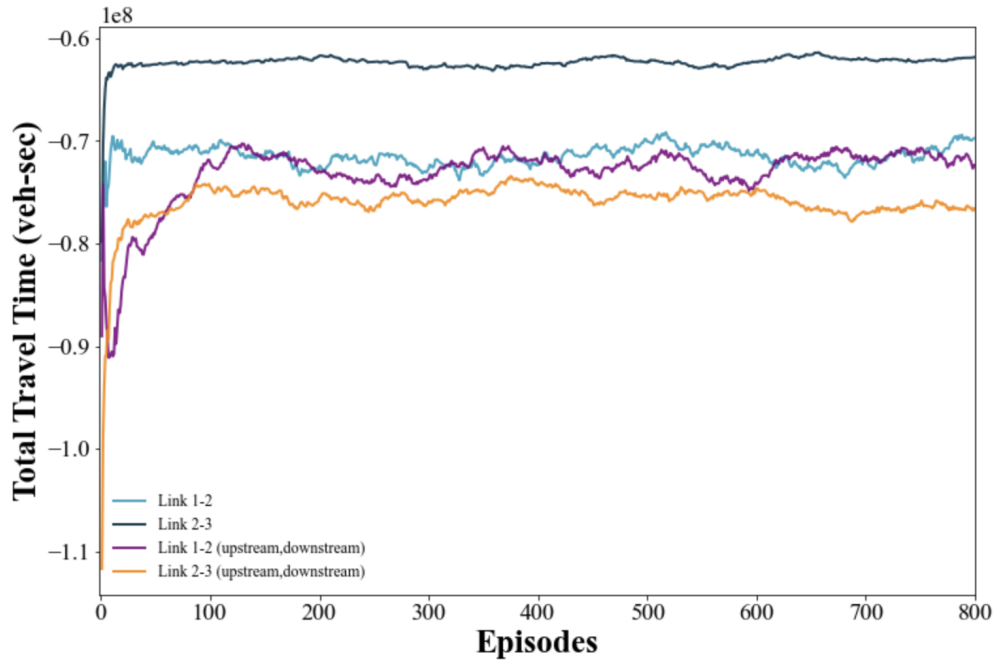


Figure 3.16: Performance of different VSLC placements to improve network mobility

speed limit provided by VSLC, it records the least number of aggressive drivers (on average 2640 drivers) and emergency situations (on average 83 times emergency brake pressed) (Figure 3.17). Considering all two parameters, this strategy improves network safety the most among all strategies while improving network mobility.

The second-best strategy is VSLC placement in the upstream and downstream of Link 1-2, with an average upstream speed limit of 60 to 65 mph and an average downstream speed limit ranging from 55 to 65 mph. As both routes have the same speed limit (40 mph), drivers take more time to decide the route to take, resulting in 58.67% higher average queue formation than the base case (233 meters on average) (Figure 3.17c). The lane change frequency is higher than the base case (on average 11.98%), contributing to 0.27% of the total delay and 5.97% of the total  $\log(TTC)$ . Additionally, the differential and abrupt nature of speed limit change in Link 1-2 leads to 797 more

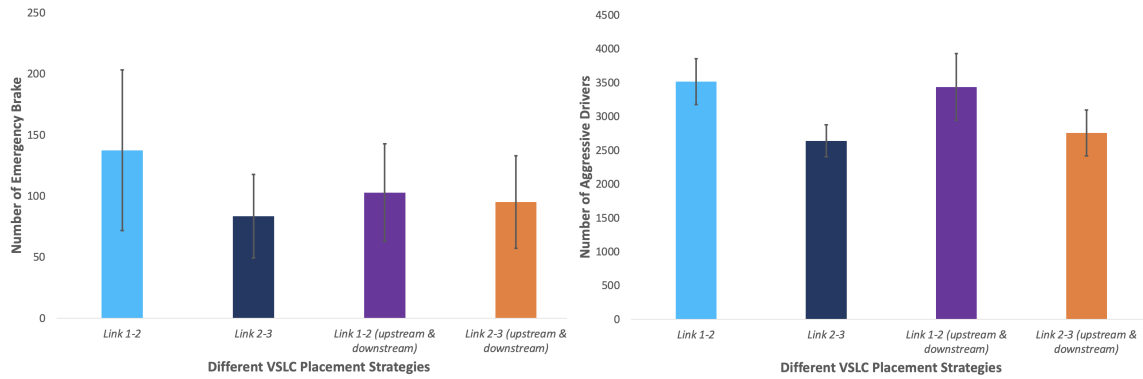
Table 3.5: Average Performance of Different VSLC Placement for Network Mobility Improvement

	Routing		Lane change frequency (%)	Network mobility		Network safety	
	Route 1 (%)	Route 2 (%)		Total Travel Time of all vehicles (sec)	Delay due to Lane change (sec)	$\sum \log(TTC)$	$\log(TTC)$ due to Lane change
Link 1-2	52.99 (2.19)	47.01 (2.19)	11.30 (2.09)	8.056e+07 (9.338e+06)	1.946e+05 (6.80e+04)	8.845e+05 (1.776e+04)	5.615e+04 (6.123e+03)
Link 2-3	65.04 (2.77)	35.96 (2.77)	11.56 (1.89)	7.166e+07 (9.951e+06)	1.283e+05 (4.933e+04)	9.274e+05 (1.348e+04)	5.679e+04 (5.302e+03)
Link 1-2 (upstream, downstream)	52.50 (2.33)	47.50 (2.33)	11.98 (5.02)	7.283e+07 (1.956e+07)	1.968e+05 (1.338e+05)	8.785e+05 (4.801e+04)	5.241e+04 (1.344e+04)
Link 2-3 (upstream, downstream)	69.91 (9.10)	30.09 (9.10)	14.13 (3.36)	8.366e+07 (1.770e+07)	2.097e+05 (8.863e+04)	9.204e+05 (1.869e+04)	6.187e+04 (7.881e+03)

aggressive drivers than the base case (Figure 3.17b). This is reflected in the higher number of vehicles pressing brakes, which is on average 103 times, higher than the base case (Figure 3.17a).

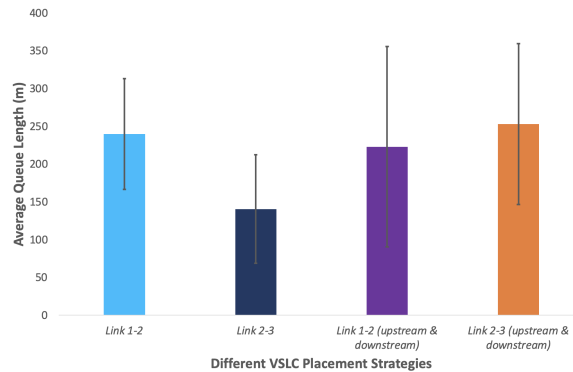
The third strategy that provides better mobility based on the subset of episodes analyzed is the implementation of VSLC in Link 1-2. VSLC placement in Link 1-2 only provides a higher speed limit, 60 to 65 mph on average. However, due to the high speed at which vehicles are traveling and vehicles' decision making to choose a route in Link 1-2, this strategy results in higher queue formation, with an average of 240 meters (Figure 3.17c), which is 70% more than the base case. As a result, the delay experienced by vehicles is 12.42% higher than the base case, with only 0.24% of this delay attributed to lane change activity. In terms of safety, the average safety reward is 4.63% lower than the base case, with 6.35% of the  $\log(TTC)$  being attributed to lane change activity (Table 3.5). Additionally, there is an average of 878 more drivers changing lanes to gain speed than in the base case, which leads to a 64% increase in the number of emergency brake situations compared to the base case (Figure 3.17).

Lastly, implementing VSLC upstream and downstream of Link 2-3 provides the same speed range (55 to 65 mph) in both areas. However, as vehicles approach the downstream area with nearly



(a) Emergency Brake Made

(b) Aggressive drivers in the network



(c) Queue Formation

Figure 3.17: Traffic behaviors for different VSLC placements for Network Mobility Improvement

the same speed, a queue forms downstream of Link 2-3. As the speed limit is higher on Route 1, most vehicles prefer using Route 1, leading to a queue forming in downstream of Link 2-3 and in Link 1-2 due to drivers choosing to use Route 1. This results in the highest average queue formation among all strategies, with an average of 253 meters or 80% more than the base case (Figure 3.17c). Consequently, this strategy leads to the highest average travel time for vehicles (Table 3.5). In terms of safety, there are an average of 117 more aggressive drivers than in the base case, resulting in a 13.8% increase in the number of emergency brake situations compared to the base case (Figure 3.17). To summarize, implementing a higher speed limit in Link 2-3 provides



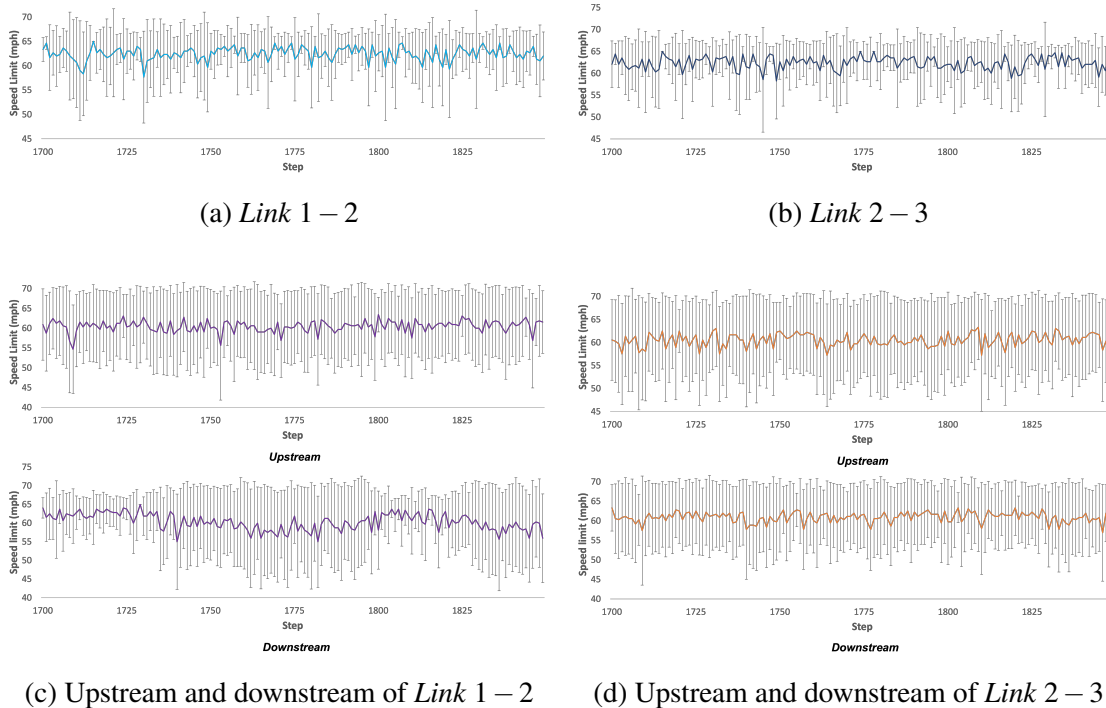


Figure 3.18: Speed profile of different VSLC implementation for improving mobility

the best network mobility improvement while forming the least bottleneck in the network and providing better network safety with fewer emergency situations. Implementing VSLC in Link 1-2 with either of the strategies results in suboptimal network mobility improvement due to the higher queue formation. Finally, implementing VSLC in the upstream and downstream of Link 2-3, rather than improving network throughput, results in additional queue formation in Link 2-3 due to the same speed range in the upstream and downstream, making it the worst-case for network mobility improvement.

### 3.7 Discussion

This study presents an approach for enhancing network mobility and safety through the implementation of VSLCs using deep reinforcement learning. The proposed DRL is built upon actor-critic architecture, namely, a twin-delayed deep deterministic policy gradient model, capable of handling continuous action spaces. In addition, the study proposes a reward function for RL models to improve network-level safety, mobility, and traffic oscillation dampening. A four-node network was used to capture the fundamental traffic events in the presence of VSLC that may not be readily apparent in a large network. The experiments show that the proposed model outperforms various DRL-based models and significantly improves network mobility and safety. Additionally, the transferability of the proposed model is assessed by varying different traffic parameters and vehicle behaviors, demonstrating its ability to adapt to changing traffic conditions. The study also presents various traffic control strategies using VSLC and highlights that Traffic engineering practice can have an adverse impact on network mobility and safety compared to controlling a routing link using VSLC. Finally, the study conducts a sensitivity analysis on the placement of VSLC in the network, identifying that controlling the speed of a routing link can significantly improve mobility and that controlling the link where vehicles make routing decisions can significantly enhance network safety. This study contributes to the field by presenting an effective approach to evaluate network safety and mobility by analyzing traffic behavior.

This research can be extended in several directions. First, a multi-agent modeling framework to improve network performance is required, which can overcome the shortcomings of a single-agent model for multiple controls. Second, studies regarding VSLC implementation on a large-scale network for traffic mobility and safety improvement should be the next step. In this case, this study might be effective in placing VSLCs in large networks. Third, the research can also be extended to VSLC implementation for different groups of vehicles in a transportation network.

# **CHAPTER 4: SPATIAL PRICING OF RIDE-SOURCING SERVICES IN A CONGESTED TRANSPORTATION NETWORK**

## **4.1 Introduction**

Ride-sourcing, pioneered by Uber in 2010, has emerged as a promising form of transportation, aiming to enhance the efficiency and reliability of point-to-point travel. Dynamic pricing, a key strategy employed by transportation network companies (TNCs), has drawn attention by dynamically adjusting fares based on supply and demand conditions. However, dynamic pricing has faced criticism for its opaque algorithms and concerns about potential exploitation of customers in an oligopolistic environment (Zha et al., 2018a). Determining optimal pricing strategies for TNCs is challenging due to the complex spatial and temporal dynamics involved, as well as the potential conflicts of interest between TNCs and society. While much research has focused on the temporal aspects of dynamic pricing, studying it from a network perspective can benefit both societal welfare and TNCs' goals. It is essential to understand the impact of TNCs' pricing decisions on system welfare to inform planning and regulation efforts. Despite ride-sourcing traffic currently having a low market share, studies have shown its negative impact on traffic congestion. Increased ride-sourcing demand has been identified as a significant contributor to congestion in cities like New York City and San Francisco. In response, some cities have implemented congestion surcharges on ride-sourcing services to alleviate congestion. While a larger fleet size reduces waiting times, it also intensifies congestion. Furthermore, the preference of young people for ride-sourcing over private vehicle ownership indicates that the market share of ride-sourcing will continue to grow, making it crucial to investigate its effects on traffic congestion.

In this dissertation, we investigate the spatial dimension of surge pricing while considering traffic

congestion during the peak period. We analyze the impacts of spatial pricing for ride-sourcing services in an interconnected transportation network using a Stackelberg framework. In the lower level, we explicitly capture the interactions between drivers' relocation, riders' mode choice, and all travelers' routing decisions. In the upper level, a single TNC provider determines spatial pricing strategies to fulfill its objective in two-sided markets. The main contribution of this work is twofold. First, we explicitly consider the relocation of drivers and transportation congestion, which not only captures the ride-sourcing behaviors more accurately but also allows for systematic analyses of the impacts of spatial pricing on transportation efficiency. Second, while optimal spatial pricing problems are challenging to solve due to a bi-level structure, we propose effective computational algorithms, with rigorous analyses on the existence, uniqueness, and convergence of the optimal solutions.

## 4.2 Literature Review

According to SAE international (2018), ride-sourcing is defined as “prearranged and on-demand transportation services for compensation in which drivers and passengers connect via digital applications”. This definition includes a broad range of mobility modes. In this paper, we focus on TNCs for automobiles.

As one of the key operational strategies, ride-sourcing pricing has attracted increasing attention from the various research communities, including transportation (e.g., (Zha et al., 2018a,b)), economics (e.g., (Bimpikis et al., 2019; Chen and Sheldon, 2016)), management science (e.g., (Guda and Subramanian, 2019)), and computer science (e.g., (Chen et al., 2015)). In addition, some studies on ride-sourcing modeling are closely related to research on taxi (e.g., (He et al., 2018; Lagos, 2000; Yang et al., 2010; Yang and Wong, 1998; Yang and Yang, 2011)) and two-sided marketplace (e.g., (Hu et al., 2021; Rochet and Tirole, 2006; Rysman, 2009)). Given the focus of this paper

is on the impacts of spatial pricing on transportation congestion, we focus on the ride-sourcing pricing issues from a transportation perspective, with references to other domains as needed.

The impacts of dynamics pricing on ride-sourcing driver supply are still inconclusive. The first conjecture, denoted as income-target theory, is that drivers quit driving once a daily income target is reached. This theory is pioneered by Camerer et al. (1997), who find negative wage elasticity of cabdriver supply in New York City. In contrast, the second conjecture, denoted as the neoclassical theory (Farber, 2005), suggests that cabdrivers are greedy for higher income, and cabdrivers supply has a positive wage rate elasticity. Farber (2005) also points out that the differences between income-target theory and neoclassical theory are in the conception and measurement of the daily wage rate. Chen and Sheldon (2016) provide empirical evidence for neoclassical theory using high-resolution UBER data. The coexistence of these two competing theories may explain the empirical evidence that the effect of dynamic pricing on UBER drivers' supply is limited (Chen et al., 2015). Zha et al. (2018a) propose a bi-level modeling framework, with lower-level equilibrium considering both income-target theory and neoclassical theory, and find that compared to static pricing, surge pricing leads to higher revenue for both the platform and drivers, but lower customer surplus during highly surged periods. The majority of these studies focus on the temporal dimension of dynamic pricing without considering spatial linkages between driver availability and traffic congestion. In this paper, due to the inconclusive impacts of dynamic pricing on driver supply, we assume that total driver supply is inelastic to spatial pricing, but drivers will respond to spatial pricing by choosing their service location.

Another school of literature investigates the potential value of dynamic pricing on ride-sourcing systems (Banerjee et al., 2015; Cachon et al., 2017; Castillo et al., 2017; Gurvich et al., 2019; Taylor, 2018). Banerjee et al. (2015) adopt a queueing-theoretical approach and find that dynamic pricing only outperforms static pricing when TNCs have imperfect knowledge of system parameters. Although network model is presented, the pricing policies only depend on the number of idle

drivers at individual locations. Castillo et al. (2017) consider the problem of “wild goose chase” when the price is set too low, and find that dynamic pricing can avoid this issue and maintain prices within a reasonable range. In contrast, in a deterministic setting, Zha et al. (2018b) adopt a geometric matching framework and find that dynamic pricing may still set price higher than the societal efficient level. However, idle vehicle relocation behaviors is not considered in Zha et al. (2018b). Karamanis et al. (2018) explore the impact of dynamic pricing for autonomous vehicles in a one-sided market for private and shared rides in both monopoly and duopoly structure. Nourinejad and Ramezani (2020) investigate TNC’s profit, driver’s wage and rider’s fare under different pricing strategies in a non-equilibrium two-sided market with pricing determined exogenously. Most of these studies focus on temporal aspects and do not consider traffic equilibrium and transportation congestion.

Bimpikis et al. (2019) are among the first to consider the spatial dimension of ride-sourcing pricing, by modeling whether and where drivers will join the platform, and where to relocate themselves when they are idle. Bimpikis et al. (2019) find that spatial pricing contributes to balance supply and demand, and more balanced demand patterns lead to higher profits and higher consumer surplus, simultaneously. However, the equilibrium formulation in (Bimpikis et al., 2019) is indeterminate and the conclusion is, therefore, best seen as an ideal bound (Zha et al., 2018b). In addition, since Bimpikis et al. (2019) mainly focus on the equilibrium of ride-sourcing services rather than traffic equilibrium, transportation topology and congestion effects are ignored. Similarly, Tang et al. (2022) examined long-term spatial pricing strategies based on long-term interaction between drivers and riders in different zones, but transportation network congestion effects are not considered. To optimize spatial-temporal pricing, Chen et al. (2021) propose a reinforcement learning approach and shows that TNC’s profit increases with dynamic pricing in real-world situations. Battifarano and Qian (2019) propose a data-driven framework to predict surge multipliers in real-time. These studies do not aim to analytically explain the interaction between different components in

ride-sourcing systems.

The congestion effect of ride-sourcing services in transportation networks is a relatively new research topic and, thus, less explored. Nie (2017) has indicated concerns about ride-sourcing services worsening traffic congestion. A more recent work by Ban et al. (2019) using a general equilibrium model to formulate the multi-agent interactions in an e-hailing system, has explicitly modeled traffic congestion under path-independent free-flow travel time assumption. However, pricing is treated as exogenous parameters and not a focus in (Ban et al., 2019). Beojone and Geroliminis (2021) investigate the impact of ride-sourcing services on traffic congestion using agent-based modeling. The study reveals larger fleet size reduces waiting time but increases congestion thereby adding more to total travel time. Vignon et al. (2021) investigated the impact of both solo and pooling ride-sourcing services on traffic congestion, which provides crucial policy insights for regulating ride-sourcing services effectively. Nevertheless, these studies do not address the impact of spatial/dynamic pricing on ride-sourcing services under congested network conditions. Li et al. (2021b) propose a network market equilibrium model to investigate the optimal spatial pricing for a ride-sourcing platform to maximize profits under different congestion pricing policies. However, the resulting model is non-convex and the global optimal is not guaranteed. We aim to fill this gap by proposing a mathematical tractable modeling framework for spatial pricing considering traffic equilibrium in a congested transportation network.

### **4.3 Methodologies**

In this section, we present mathematical models for ride-sourcing systems in two cases: (1) ignoring waiting time (Section 4.3.1), and (2) considering waiting time (Section 4.3.2). The waiting time used in this paper is defined as follows. For a rider, waiting time is the duration between a request being made and the rider being picked up. For a driver, waiting time is the duration

between the time when the driver arrives at the desired service location waiting for ride requests and the time when the driver picks up a rider. Based on this definition, the waiting time includes matching time and pick up time. This definition is consistent with existing ride-sourcing literature (e.g., (Yang and Yang, 2011; Zha et al., 2016)).

#### 4.3.1 Mathematical Modeling: Ignoring Waiting Time

Firstly, we ignore the impacts of waiting time on the behaviors of drivers and riders. This assumption will be relaxed in Section 4.3.2. Models ignoring waiting time are appropriate when waiting time is an insignificant factor for drivers' and riders' decision making. For example, for long-term transportation planning problems with congestion, relocation/travel time can significantly exceed waiting time. In addition, travelers may be able to request rides earlier than their intended departure time, so that ride request decisions are less sensitive to waiting time.

##### 4.3.1.1 Traffic Modeling

Denote a transportation network by a directed graph  $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ , where  $\mathcal{N}$  is the set of nodes (indexed by  $n$ ) and  $\mathcal{A}$  is the set of links (indexed by  $a$ ). The service area we consider is a metropolitan area where ride-sourcing demand and supply are aggregated in predefined zones (e.g., census tracts or traffic analysis zones), represented by nodes. A link is a sequence of roads that connects two nodes. Given driver node set  $\mathcal{R} (\subset \mathcal{N})$  and rider node set  $\mathcal{S} (\subset \mathcal{N})$ , drivers will relocate from  $r (\in \mathcal{R})$  to  $s (\in \mathcal{S})$  to provide ride-sourcing services depends on spatial pricing. In other words, there are two "origins" used in this paper. A driver's origin, denoted as  $r$ , is the driver's current location before he/she makes relocation; a rider's origin or a driver's relocation destination, denoted as  $s$ , is where the rider makes a ride request or the driver relocates to. If a driver decides to wait at his/her current location or has accepted a trip in the same area,  $r$  equals  $s$ .



Drivers are more likely to relocate to locations with higher expected service prices. To capture this effect, we propose utility function (4.1). We assume that driver relocation decisions follow multinomial logistic models, with explanatory variables of locational attractiveness  $\beta_{0,s}$ , relocation time  $t_{rs}$ , and locational prices  $\rho_s$ . Locational attractiveness indicates how a location is favored by drivers because of factors other than time and price, such as safety, rider density, and easiness of getting returning trips. Locational prices  $\rho_s$  represent the expected trip standard rates originated from  $s$  times origin-base surge multipliers determined by TNC. The modeling framework is flexible to incorporate more explanatory factors. The deterministic components of drivers' utility function are given in (4.1).

$$U_{rs} = \beta_{0,s} - \beta_1 t_{rs} + \beta_2 \rho_s \quad (4.1)$$

where:

- $U_{rs}$  : deterministic component of utility measure for a driver going from  $r$  to  $s$ ;
- $\beta$  : utility function parameters (model input);
- $t_{rs}$  : equilibrium travel time from  $r$  to  $s$ ;
- $\rho_s$  : locational price at  $s$ .

Utility function (4.1) describes the utility of drivers at location  $r$  when they make relocation decisions from  $r$  to  $s$ . Utility function (4.1) does not aim to capture the heterogeneity of drivers at location  $r$  nor reflect how much each driver will actually get paid, since supply and demand may not be balanced at each location and drivers may be matched with trips with different trip lengths. But when drivers make relocation decisions, they do not know the length of trips that they are going to be matched with by the TNC. In other words, drivers' revenue from the next trip is a random variable before they relocate. The expectation of drivers' revenue depends on the expectation of trip length and the origin-based surge multipliers of their relocation destinations. Therefore, for

drivers, the relocation decision making is independent of the specific destination of the next trip. Instead, it will depend on the expected trip fares originated from the region a driver relocates to.

While drivers have the flexibility to relocate themselves to other locations, travelers typically have predefined travel origins and destinations. In this study, we assume the travel demand  $D_s$  is endogenously given but they can choose the travel modes between requesting TNC services and driving. At each location  $s (\in \mathcal{S})$ , we start with a linear ride-sourcing demand function, as defined in (4.2), to calculate ride-sourcing demand  $d_s$ . Driving demand is  $D_s - d_s$ . We note that the linear demand function (4.2) is used to keep a clear focus of this paper and it does not restrict the applicability of our modeling and computational approach for more sophisticated demand models (see B.3.1).

$$d_s = D_s - b_s \rho_s, \rho_s \leq D_s / b_s \quad (4.2)$$

where:

$d_s$  : flow rate of riders requesting ride-sourcing service at location  $s$ ;

$D_s, b_s$  : total travel demand originated from  $s$  and slope of demand function (model input).

To calculate the ride-sourcing OD demand  $d_{sk}$ , we need to multiply distribution coefficient  $\delta_{sk}$  with  $d_s$ .  $d_s$  is independent of destinations because it is the ride-sourcing demand aggregating all possible destinations.  $\delta_{sk}$  will be influenced by destination characteristics, trip lengths, and other related parameters. This modeling strategy is widely adopted in the classic “four-step travel demand model” for transportation planning (McNally, 2007), in which travelers are modeled as making four distinct decisions: trip generation, destination choice, modal choice, and route choice. To facilitate modeling and computation, these four steps are typically treated separately. In this study, we focus on modeling how travelers choose between different travel modes and routes and consider the total trip generation  $D_s$  and trip destination choice  $\delta_{sk}$  as exogenous parameters. Saying that, we

note that traffic distribution factors can be endogenously determined within the proposed modeling framework based on the travel time ( $t_{sk}$ ) and travel distance ( $l_{sk}$ ) from  $s$  to  $k$ , as well as locational attractiveness at location  $k$ . The reason travel time/distance is a good proxy to spatial pricing is because ride-sourcing trip fares are typically determined by trip time and distance<sup>1</sup>. The longer the travel time/distance is for a destination  $k$ , the higher the prices riders need to pay. Therefore, riders will less likely to choose that destination. This effect can be directly modeled using combined distribution and assignment (CDA) model with utility function (4.3) of destination choice for riders, where  $b_0, b_1, b_2$  are utility coefficients corresponding to locational attractiveness, travel time, and trip distance.

$$U_{sk} = b_{0,k} - b_1 t_{sk} - b_2 l_{sk} \quad (4.3)$$

The interactions between drivers, riders, and other passenger vehicles will endogenously determine the mobility of the transportation system. We extend the combined distribution and assignment (CDA) model (Wilson, 1969) to describe the interactions between drivers' relocation and the routing of all ride-sourcing vehicles and conventional passenger vehicles. The model is described in (4.4).

---

<sup>1</sup><https://help.uber.com/riders/article/how-are-fares-calculated/?nodeId=d2d43bbc-f4bb-4882-b8bb-4bd8acf03a9d>

$$\underset{\hat{v}, \check{v}, q \in +}{\text{minimize}} \quad \sum_{a \in \mathcal{A}} \int_0^{v_a} t_a(v_a) du + \frac{1}{\beta_1} \sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{S}} q_{rs} (\ln q_{rs} - 1 - \beta_2 \rho_s - \beta_{0,s}) \quad (4.4a)$$

subject to

$$v = \sum_{r \in \mathcal{R}, s \in \mathcal{S}} \hat{v}_{rs} + \sum_{s \in \mathcal{S}, k \in \mathcal{H}} \check{v}_{sk}, \quad (4.4b)$$

$$(\hat{\lambda}_{rs}) \quad A \hat{v}_{rs} = q_{rs} E_{rs}, \quad \forall r, s \quad (4.4c)$$

$$(\check{\lambda}_{sk}) \quad A \check{v}_{sk} = \delta_{sk} D_s E_{sk}, \quad \forall s, k \quad (4.4d)$$

$$(\gamma_r) \quad \sum_{s \in \mathcal{S}} q_{rs} = Q_r, \quad \forall r \quad (4.4e)$$

where:

$\hat{v}_{rs}, \check{v}_{sk}$  : link flow vector (each row corresponds to a link,  $a$ ) grouped by OD for drivers and passenger vehicles, respectively;

$q$  : drivers' relocation flow vector (each row corresponds to an OD,  $rs$ );

$v_a$  : aggregate traffic flow on link  $a$ ;

$t_a(\cdot)$  : travel time function of link  $a$ ;

$A$  : node-link incidence matrix of network, with 1 at starting node and  $-1$  at ending node;

$E_{rs}$  : O-D incidence vector of O-D pair  $rs$  with 1 at origin  $r$ ,  $-1$  at destination  $s$ . If  $r = s$ ,

$$E_{rs} = 0;$$

$\delta_{sk}$  : traffic distribution factor from  $s$  to  $k$  (model input);

$Q_r$  : drivers initial availability at  $r$ ;

$\check{\lambda}, \hat{\lambda}, \gamma, \eta$  : dual variables of the corresponding constraints.

In the objective function (4.4a), the first term corresponds to the travel cost as modeled in a classic static user equilibrium model (Beckmann et al., 1956), the second term involving  $q^{rs} \ln q^{rs}$  corresponds to the entropy of trip distribution, and the remaining terms are constructed in such a way

that the first order derivative of the Lagrangian function with respect to  $q_{rs}$  resembles the utility measure (4.1) of drivers. This objective function does not have an intuitive physical interpretation to the authors' best knowledge, but it guarantees the first Wardrop principle (Wardrop, 1952b) and the multinomial logit location choice assumption being satisfied, as formally stated in **Lemma 4.3.1**.

**Lemma 4.3.1.** *(combined distribution and assignment for ride-sourcing system) The optimal solutions  $(\hat{v}^*, \check{v}^*, q^*)$  of problem (4.4) are the equilibrium solutions for Wardrop user equilibrium and the relocation choice of drivers with multinomial logit model (4.1) given spatial prices  $\rho$ .*

Proof. See B.1.

Constraint (4.4b) calculates the aggregate link flow  $v_a$  from the link flow associated with OD pairs  $rs$  and  $sk$ . Constraints (4.4c, 4.4d) are link-based formulations of flow conservation at each node for each OD pair. Note that in the flow conservation constraint (4.4d), we consider not only trips fulfilled by ride-sourcing services ( $d_s$ ), but also trips fulfilled by driving  $D_s - d_s$ . The total traffic originating from node  $s$  will be  $D_s$ . The reason is that ride-sourcing traffic flow is (currently) a small proportion of the link flow and considering external traffic is important to estimate traffic congestion. We model available driver per unit time period at location  $r$  as flow rate  $Q_r$  as given, which will relocate to different destinations  $s$  with a flow rate  $q_{rs}$ , as specified in Constraint (4.4e). One can also extend model (4.4) to consider the network effects of rider's destination on the spatial distribution of available drivers, i.e., if a driver takes a rider from  $s$  to  $r$ , then thereafter the driver will be at  $r$  (see B.3.2). The equilibrium path travel time  $t_{rs}$  is reflected by  $\hat{\lambda}_{rs}^{a_i} - \hat{\lambda}_{rs}^{a_j}$ . For details, please see the proof of Lemma 4.3.1.

Note that there are two types of ride-sourcing traffic in model (4.4). The first type is due to relocation or deadheading trips (i.e. no passengers on board); the second type is after ride-sourcing vehicles pick up passengers, drivers will drive riders to the riders' destinations, in which case ori-

gin and destination are known. The first type of ride-sourcing trip is  $Q_r$  in constraint (4.4e), which does not have specific destinations.  $D_s$  includes the second type of ride-sourcing trips and other driving trips. We do not need to separate travel demand by ride-sourcing and driving in mode (4.4) because the OD travel demand of vehicle traffic from  $s$  to  $k$  is fixed, which typically can be estimated from classic travel demand models. Although rider demand  $d_s$  does not directly enter into the model (4.4),  $d_s$  is endogenously determined and implicitly influences the locational prices  $\rho_s$  when we solve the multi-agent optimization models simultaneously.

Our proposed model can be extended to model elastic drivers' supply  $Q_r$  as a function of prices using discrete choice models. To do that, one can add a dummy relocation destination node for drivers in the transportation network to represent that drivers may decide to quit the service for the studied time period due to extremely low spatial prices offered at the candidate relocation destinations. The price parameter at the added dummy node reflects the benchmark revenue drivers may earn or expect to earn using their time. This approach does not need to modify our modeling (e.g., model (4.4)) and computational strategies.

#### 4.3.1.2 TNC Behaviors

Due to the uncoordinated behaviors of a large number of drivers and riders, individual drivers and riders do not have market power and only respond to the spatial prices set by the TNC. The pricing decisions of a TNC will influence the traffic equilibrium solutions of (4.4) and rider demand (4.2), which forms a mathematical programming with equilibrium constraints (MPEC) problem.

In this paper, we only consider TNC using origin-based pricing to influence the spatial distribution of ride-sourcing demand and supply. Origin-based pricing is widely adopted by TNCs (Zuniga-Garcia et al., 2020). The trip origin determines the amount of surge pricing for a trip, instead of the driver's location or trip destination. When prices are surging, TNCs will charge riders in the surge

zone a multiplier to standard rates. The income of drivers serving the trips originated from the surge zone also scales up by the surge multiplier. The origin-based prices  $\rho_s$  used in this manuscript can be interpreted as regular fares for expected trip length multiplied by origin-base surge multipliers. Note that as long as the expected prices paid to drivers and paid by riders are linearly related to  $\rho_s$ , distinguishing payments from/to riders/drivers does not impact the equilibrium outcomes. This is because the impact will be implicitly captured by the coefficients  $\beta_2$  in the utility function (4.1) and  $b_s$  in the demand function (4.2).

The TNC decision-making can be modeled as a bi-level problem (4.5). We assume that only one TNC in our study area, who would like to minimize the imbalance between supply and demand (see, objective (4.5a))<sup>2</sup>, which has the advantage of improving driver/rider satisfaction and ultimately leads to long-term sustainable development and profits. A comparison between different objectives (minimizing imbalance versus maximizing short-term profits) will be presented in Section 4.5. Constraint (4.5b) calculates the imbalance as the difference between the flow of riders requesting services  $d_s$  and the number of drivers arriving at this location  $\sum_{r \in \mathcal{R}} q_{rs}$ . Calculation of traffic equilibrium patterns  $q_{rs}$  and ride-sourcing demand  $d_s$  are included in constraint (4.5c).

$$\begin{array}{ll} \text{minimize} & \sum_{s \in \mathcal{S}} m_s \\ \rho \in \mathcal{S} & \end{array} \quad (4.5a)$$

subject to

$$m_s = \left| \sum_{r \in \mathcal{R}} q_{rs} - d_s \right|, \forall s \quad (4.5b)$$

$$(4.4), (4.2) \quad (4.5c)$$

where:

---

<sup>2</sup>Balancing supply and demand is the goal of dynamic pricing proposed by all TNC companies. See, for example, <https://www.ridester.com/training/lessons/surge-primetime-boost-primezones/> (visited on Feb. 26, 2020).

$m^s$  : demand-supply imbalance at  $s$ .

While constraint (4.5b) can be reformulated as linear constraints, the bi-level structure still makes the whole problem challenging to solve to global optimal (Sinha et al., 2017; Yang and H. Bell, 1998). In Section 4.4, we propose an equivalent single-level convex reformulation of (4.5), which can be efficiently solved by commercial nonlinear solvers (e.g. IPOPT (Wächter, 2009)).

#### 4.3.2 Mathematical Modeling: Considering Waiting Time

In this section, we explicitly model the impacts of waiting time. Waiting time could influence the relocation decisions of drivers and the mode choices of riders. From an individual perspective, waiting time is heterogeneous and uncertain for each driver and rider. In this paper, we focus on the aggregated behaviors of drivers/riders and use expected waiting time to reflect the time costs to provide or receive services in a zone on average.

In reality, there are two phases, “matching” and “pick-up” after a ride request is made. For transportation system modeling purposes, Cobb-Douglas-type aggregate matching functions are widely adopted (Yang and Yang, 2011; Zha et al., 2016) to estimate the “meeting” rate and total waiting time (Li et al., 2021c). This modeling strategy is validated in Zha et al. (2016) using an agent-based simulation study. To consider different matching technologies (e.g., matching radius and matching time interval), a general matching function has been proposed for both single-interval optimization (Yang et al., 2020) and stationary analyses (Li et al., 2021c), which includes Cobb-Douglas-type matching functions as a special case. The main disadvantage of these studies is that there may not be a closed-form solution for the matching function and the waiting time could only be implicitly solved from a system of equations. Therefore, in this paper, we adopt a Cobb-Douglas matching function (see Equation (4.6)) as a starting point to describe the total waiting time of drivers and



riders, which includes waiting time in the matching phase and pick-up phase.

A Cobb-Douglas function is used mainly because of three reasons. First, the Cobb-Douglas function has a constant elasticity, which is specified by parameters  $\alpha_1$  and  $\alpha_2$ . The elasticity can reflect the efficiency of the matching technology adopted by a ride-sourcing platform. Second,  $\alpha_0$  has a nice interpretation, which is related to the area of the matching zone and the cruising speed of vacant vehicles (Yang and Yang, 2011). Third,  $\alpha_1 + \alpha_2$  is larger than, equal to, or smaller than one when the matching function is increasing, constant, or decreasing returns to scale. For a given area, it is typically assumed that  $\alpha_1 + \alpha_2 > 1$  to represent increasing returns to scale of matching rate, because the increasing density of supply and/or demand will make it easier to match drivers with riders. Note that  $N_D$  and  $N_R$  are the number (instead of flow rate) of available drivers and awaited riders in a zone, which may change over our study time period due to a potential imbalance of supply and demand.

$$m(t) = \alpha_0 N_D^{\alpha_1}(t) N_R^{\alpha_2}(t) \quad (4.6)$$

Denote the locational average and total waiting time for drivers ( $D$ ) and riders ( $R$ ) as  $t_D$ ,  $T_D$ ,  $t_R$ , and  $T_R$ , respectively. Based on matching rate function (4.6), average waiting time for drivers  $D$  and riders  $R$  can be calculated in Equations (4.7)<sup>3</sup>, where  $f_i$  ( $i \in \{D, R\}$ ) is the flow rate of drivers/riders (i.e.,  $f_{D,s} = \sum_{r \in R} q_{rs}$ ,  $f_{R,s} = d_s$ ) and  $T$  is the study period. Note that (4.7) is location specific, but we omit the location index  $s$  for all of the variables in (4.7) for notation conciseness.

---

<sup>3</sup>Equation (4.7) is explained graphically in B.2

$$\dot{N}_i(t) = f_i - m(t) = f_i - \alpha_0 N_D^{\alpha_1}(t) N_R^{\alpha_2}(t), \quad (4.7a)$$

$$T_i = \int_0^T N_i(t) dt, \quad \forall i \in \{D, R\} \quad (4.7b)$$

$$t_i = \frac{T_i}{f_i T} = \frac{\int_0^T N_i(t) dt}{f_i T}, \quad (4.7c)$$

Equation (4.7a) calculates the change rate of  $N_D$  (i.e.  $\dot{N}_D$ ) as how many drivers relocate to  $s$  (i.e.,  $\sum_{r \in R} q_{rs}$ ) minus matching rate  $m$ . Similarly, the changing rate of  $N_R$  (i.e.,  $\dot{N}_R$ ) is determined by rate of rider demand in  $s$  (i.e.,  $d_s$ ) minus matching rate  $m$ . Equations (4.7b) and (4.7c) calculate the total and average waiting time for drivers and riders over  $T$ , respectively.

In (Yang and Yang, 2011) and some other subsequent studies, the authors assume the number of available drivers and riders are stationary, which implies that the matching rate will always equal to drivers/riders arrival rate. In contrast, we only assume that the *arrival rates* of drivers and riders are stationary during our studied time period. In other words, the flow rate could be greater, equal, or even lower than the matching rate if there is a reasonable level of available drivers and riders at the beginning of the studied period. When the flow rate is higher (or lower) than the matching rate for a certain duration, the number of available drivers and riders will increase (or decrease) and the matching rate will increase (or decrease) based on equation (4.6). The reason for this relaxation is to have a more realistic representation of the real ride-sourcing market, where the matching rate is typically not constant and available drivers and riders may fluctuate during the studied period. Our proposed model can be easily degenerated to consider the case when matching rate equals to drivers/riders flow rate by specifying  $f_i = m(t), \forall i \in \{D, R\}$  in equation (4.7a) so that  $\dot{N}_i(t) = 0$ .

While solving Equation (4.7) analytically is challenging,  $t_i (\forall i \in \{D, R\})$  can be solved numerically given  $f_D, f_R$  and boundary conditions  $N_D(0), N_R(0)$ . We then apply linear regression techniques

to estimate the relationship between  $t_i$  and  $f_i$ ,  $\forall i \in \{D, R\}$ . The reason we need to use regression techniques is that the relationship between  $t_i$  and  $f_i$  is not well documented in the literature. We select  $\alpha_0 = 0.1, \alpha_1 = \alpha_2 = 0.6$  for our numerical simulation to generate a sequence of  $t_i$  and  $f_i$  following system dynamic equations (4.7). With these data generated, a linear regression model is used to estimate the coefficients in (4.8). When  $\alpha_0 = 0.1, \alpha_1 = \alpha_2 = 0.6$ , we have  $a_0 = 6.29, a_1 = 2.24, a_2 = -2.40$  as least square estimators.  $|a_2| > |a_1|$  implies that if both drivers and riders double their flow to a region, the waiting time for drivers and riders on average will be reduced (i.e., increasing return of scale for matching rate). The linear regression data generation procedures and output summary are presented in B.2.

$$t_i = a_0 f_i^{\alpha_1} f_{-i}^{\alpha_2} \quad \forall i \in \{D, R\} \quad (4.8)$$

**Remark:** Our modeling framework can have an equation similar to Equation (4.6) to explicitly model the pickup time at a given zone, i.e.,  $t^{pickup}(t) = \alpha_0^{pickup} N_D^{\alpha_1^{pickup}}(t) N_R^{\alpha_2^{pickup}}(t)$ . Since the pickup time could change over time, we can estimate the average pickup time over the studied time period as  $\bar{t}^{pickup} = \frac{\int_0^T t^{pickup}(t) dt}{T}$ . In this case, the total waiting time ( $t_i^{total}$ ) the drivers or the riders spend from placing ride requests to being picked up equates to  $\bar{t}^{pickup} + t_i$ . Following the same strategies to estimate the matching time  $t_i$  (i.e., Equations (4.7) and (4.8)), we can estimate the relationship between the average pickup time  $\bar{t}^{pickup}$  and drivers/riders flow  $f_R$  and  $f_D$  using regression techniques. Note that this extension does not affect the proposed computational strategies below.

To incorporate waiting time in a unified CDA framework, we define an augmented transportation network,  $\bar{\mathcal{G}} = (\bar{\mathcal{N}}, \bar{\mathcal{A}})$ . The augmenting procedures from  $\mathcal{G}$  to  $\bar{\mathcal{G}}$  is illustrated in Figure 4.1 using a three-node transportation network, which has one driver node (node 1) and two rider nodes (node 2 and 3). For each rider node  $s$  in Figure 4.1a, we created  $s, s',$  and  $s''$ , representing nodes for

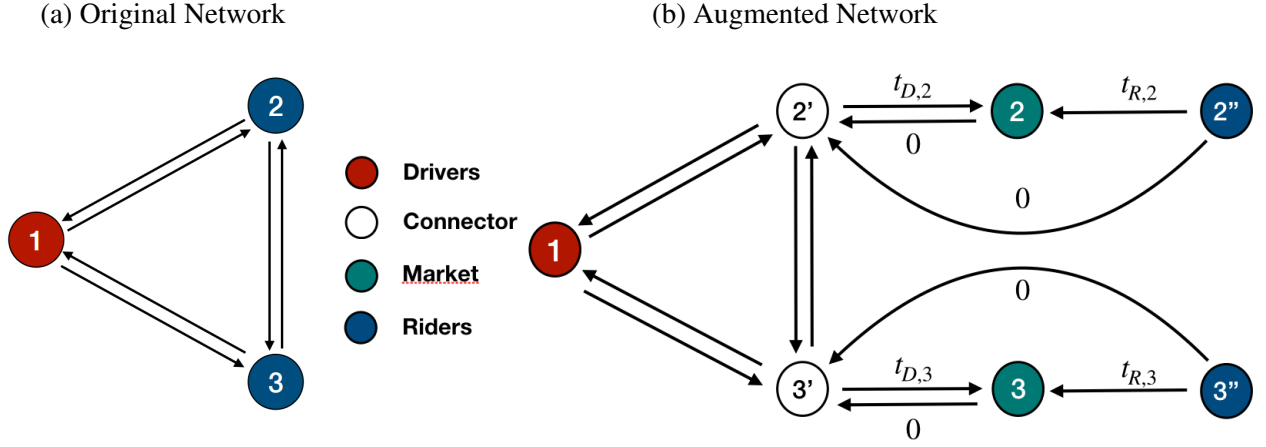


Figure 4.1: Illustration of Network Augmentation

markets, connectors, riders, respectively, with directed links connecting between them as shown in Figure 4.1b<sup>4</sup>. In augmented network  $\bar{\mathcal{G}}$ , link flow  $f_{s's}$ ,  $f_{s''s}$  and  $f_{s''s'}$  represent the flow of driver supply, rider demand, and travelers choosing to drive at node  $s$ , while the link cost  $t_{s's}$  and  $t_{s''s}$  are the waiting time for drivers ( $t_{D,s}$ ) and riders ( $t_{R,s}$ ), respectively. If travelers choose to drive, there is no waiting time, so that link cost  $t_{s''s'} = 0$ . In this paper, we do not consider cross-node matching. This is because we assume that each node represents a relatively large area (e.g., zipcode/city) where the majority of matching is within the node. However, the augmented network (Figure 4.1) can be further extended to consider the cross-node matching by adding single directional arrow from  $2'$  to  $3$  and from  $3'$  to  $2$ . The “travel time” on these links represents the cross-node waiting time.

With an augmented network  $\bar{\mathcal{G}}$ , the utility function of drivers is identical with (4.1). Note that  $t_{rs}$  in (4.1) will include not only the travel time from  $r$  to  $s'$ ,  $t_{rs'}$ , but also the waiting time at  $s$ ,  $t_{s's}$ . The drivers relocation and routing behaviors, therefore, can be modeled as (4.4) (replacing  $\mathcal{G}$  with  $\bar{\mathcal{G}}$ ), given spatial pricing  $\rho$ .

<sup>4</sup>From Figure 4.1, we can see that ignoring waiting time is a special case of considering waiting time when  $a_0 = 0$  in (4.8), which makes link cost  $t_{D,2}, t_{R,2}, t_{D,3}, t_{R,3} = 0$  in Figure 4.1b. This makes Figure 4.1b identical as Figure 4.1a.

For travelers at location  $s$ , the ride requesting demand function is not only dependent on the prices  $\rho_s$ , but also depends on the waiting time,  $t_{s''s}$ . We formulate the choice of ride-sourcing over driving using a binary logit model, with ride-sourcing utility function (4.9),

$$U_{s''s} = \beta'_{0,s} - \beta'_1 t_{s''s} - \beta'_2 \rho_s \quad (4.9)$$

where  $\beta'_{0,s}, \beta'_1, \beta'_2$  are the utility coefficients for ride-sourcing attractiveness, waiting time, and ride-sourcing prices. Since the on-road travel time for driving and ride-sourcing are the same, travel time will be canceled out when travelers make mode choices and do not need to be included in (4.9). Note that the rider demand couples with waiting time (see (4.8) and (4.9)), we will formulate the travelers' mode choices using the CDA model in the augmented network  $\bar{\mathcal{G}}$ , where the destination  $s$  represents travelers choosing ride-sourcing services, and destination  $s'$  represents travelers choosing driving. The CDA model for riders' mode choice at each location  $s$  is shown in (4.10), where the first term corresponds to the conventional Wardrop user equilibrium, and the second term reflects the binomial logit choice between ride-sourcing and driving.

$$\underset{f_{R,s} \in +}{\text{minimize}} \quad \int_0^{f_{R,s}} t_{R,s}(f_{D,s}, u) du + \frac{1}{\beta'_1} [f_{R,s} (\ln f_{R,s} - 1 + \beta'_2 \rho_s - \beta'_{0,s}) + (D_s - f_{R,s}) (\ln(D_s - f_{R,s}) - 1)] \quad (4.10)$$

Similar to formulation (4.5) for the “ignoring waiting time” case, the TNC decision making considering waiting time can be modeled as a bilevel problem, as shown in (4.11), where objective (4.5a) minimizes total demand-supply imbalance (calculated by (4.5b)), subject to drivers, riders, and all other travelers user equilibrium (4.4) and travelers' mode choices (4.10) in the augmented network  $\bar{\mathcal{G}}$ .

$$\text{minimize}_{\rho \in \mathcal{S}} \quad (4.5a) \quad (4.11a)$$

subject to

$$(4.5b), (4.4, \text{replacing } \mathcal{G} \text{ with } \bar{\mathcal{G}}), (4.10) \quad (4.11b)$$

**Remark:** In this paper, we assume that travelers only choose between ride-sourcing and driving to keep a clear focus on the spatial pricing issues in ride-sourcing systems. However, we note that the proposed modeling framework can incorporate more sophisticated mode choice models. This can be seen from models (4.9) and (4.10), where riders choosing ride-sourcing or not (binary logit model) can be reformulated as a combined distribution and assignment (CDA) model. Since the CDA model can naturally consider more than two destinations (Florian et al., 1975), if other transportation modes are of interest, we can introduce dummy destination nodes in Figure 4.1 to represent other travel modes and extend model (4.10) to solve for a multinomial logit model.

## 4.4 Solution Approach

### 4.4.1 Solution Approach: Ignoring Waiting Time

We start by showing that when problem (4.5) is solved to optimal, the ride-sourcing market will be balanced at each location. In addition, optimal solutions  $\rho^*$  is unique. This is more formally stated in Lemma 4.4.1.

**Lemma 4.4.1.** *(balancing ride supply and demand)  $\rho^*$  optimizes problem (4.5) if and only if ride supply and demand are balanced at each location given  $\rho^*$ . In addition,  $\rho^*$  is unique and  $\rho_s^* \in [(D_s - \sum_{r \in R} Q_r)/b_s, D_s/b_s], \forall s \in S$ .*

Proof. See B.1.

Lemma 4.4.1 says that problem (4.5) is equivalent to the problem of finding the unique  $\rho^*$  that balances supply and demand at each location. Notice that  $\rho^*$  could be negative under extreme scenarios when too many available drivers exist in the system compared to rider demand. In other words, TNC may provide incentives for riders to use ride-sourcing services only when  $D_s < \sum_{r \in R} Q_r$ . However, this will not happen in reality given reasonable driver supply  $Q_r$ , since drivers will dynamically respond to extremely low prices (Bimpikis et al., 2019). On the other hand, based on Lemma 4.4.1, since  $\rho_s^* \leq D_s/b_s$ ,  $d_s^* \geq 0$ . In other words,  $d_s$  for all  $s$  cannot be negative in the optimal solution.

Leveraging Lemma 4.4.1, the optimal spatial prices can be obtained by solving a single-level convex optimization problem, which will be illustrated in the remainder of this section.

Problem (4.5) has three types of agents: drivers, riders, and TNC. Because of Lemma 4.4.1, the decision making of TNC can be reformulated as finding  $\rho^*$  such that

$$\sum_{r \in \mathcal{R}} q_{rs}^*(\rho^*) = D_s - b_s \rho_s^*, \forall s \in \mathcal{S} \quad (4.12)$$

where  $q_{rs}^*(\rho^*)$  is the optimal solutions of (4.4) given  $\rho^*$ . This problem falls within the framework of multi-agent optimization problem with equilibrium constraints (MOPEC) (Ferris and Wets, 2013), where drivers make relocation decisions; riders decide whether to request ride-sourcing services; all travelers make route choices decisions; and the system is subject to equilibrium conditions (4.12). MOPEC has been applied in transportation, energy, economics domains, such as charging infrastructure planning (Guo et al., 2016b), hydro-thermal electricity systems optimization (Philpott et al., 2016), renewable energy supply chain planning (Guo and Fan, 2017), transportation and power system interdependence (Baghali and Guo, 2021; Guo et al., 2021), and computational Walrasian equilibrium (Deride et al., 2019).

Define two types of dummy agents in our problem, *Driver* and *Rider*. Given  $\rho$ , *Driver* solves problem (4.4), which determines the drivers relocation decision from  $r$  to  $s$ ,  $q_{rs}(\rho), \forall r, s$ ; and *Rider* solves problem (4.13), which determines the riders demand  $d_s(\rho)$  at each location  $s$ . Note that problem (4.13) is a constructed unconstrained quadratic programming problem, which has a closed-form solution  $d_s^*(\rho) = D_s - b_s \rho_s, \forall s \in \mathcal{S}$ , identical to the demand function (4.2) we assumed.

$$\text{minimize}_{d \in \mathbb{S}_+^{\mathcal{S}}} \sum_{s \in \mathcal{S}} \frac{1}{b_s} [-D_s d_s + b_s \rho_s d_s + \frac{1}{2} (d_s)^2] \quad (4.13)$$

The equilibrium constraints (4.12) can be reformulated as (4.14).

$$\sum_{r \in \mathcal{R}} q_{rs}^*(\rho) = d_s^*(\rho), \forall s \in \mathcal{S} \quad (4.14)$$

We construct a single-level convex optimization problem (4.15), which will yield the same spatial pricing decisions as bilevel formulation, as stated in Theorem 4.4.2.

$$\text{minimize}_{\hat{v}, \check{v}, q, d \in \mathbb{S}_+} \frac{\beta_1}{\beta_2} \sum_{a \in \mathcal{A}} \int_0^{v_a} t_a(v_a) du + \frac{1}{\beta_2} \sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{S}} q_{rs} (\ln q_{rs} - 1 - \beta_{0,s}) + \sum_{s \in \mathcal{S}} \frac{1}{b_s} (\frac{d_s^2}{2} - D_s d_s) \quad (4.15a)$$

subject to

$$(4.4b \sim 4.4e)$$

$$(\rho_s) \quad \sum_{r \in \mathcal{R}} q_{rs} = d_s, \forall s \quad (4.15b)$$

The objective function (4.15a) is a combination of scaled objective functions (4.4a) and (4.13) without  $\sum_{r \in \mathcal{R}} \rho_s q_{rs}$  and  $\rho_s d_s$  terms. The purpose of scaling (4.4a) and (4.13) is to convert the units



of both objective functions into monetary units so that they are addable. In addition, the optimal dual variables  $\rho_s^*$  of constraints (4.15b) will have a monetary unit in dollars and balance supply and demand at each location. Therefore,  $\rho^*$  can be interpreted as spatial prices to solve problem (4.5) based on Lemma 4.4.1.

**Theorem 4.4.2.** *(single-level convex reformulation)  $\rho$  solves bi-level problem (4.5) if and only if  $\rho$  is the optimal dual variable corresponding to constraint (4.15b) in single-level problem (4.15).*

Proof. See B.1.

Problem (4.15) is a non-linear convex problem, with a convex objective function and linear constraints. So it can be efficiently solved to global optimal by commercial nonlinear solvers, such as IPOPT. For extreme large-scale problems, this reformulation opens up opportunities to apply classic and efficient transportation network solution algorithms, such as the Frank-Wolfe algorithm (Sheffi, 1985) and Evans' procedure (Evans, 1976). In this paper, we focus on small/medium transportation networks to draw policy insights and leave the development of advanced algorithms for extremely large networks for the future.

#### 4.4.2 Solution Approach: Considering Waiting Time

Directly solving TNC upper-level problem (4.11) is challenging due to non-convexity. We restate (4.11) as a problem of finding a maxinf-point for a bivariate function (bi-function)  $W(\rho, \phi)$ , as shown in (4.16).

$$W(\rho, \phi) = - \sum_{s \in \mathcal{L}} \phi_s \text{ES}_s(\rho), \quad \text{on } \mathcal{S} \times \Delta_S \quad (4.16)$$

where  $\Delta_S$  corresponds to the  $S$ -dimensional unit simplex.  $\text{ES}_s$  is the square of demand-supply

imbalance, defined in (4.17).

$$\text{ES}_s(\rho) = \left[ \sum_{r \in \mathcal{R}} q_{rs}^*(\rho) - f_{R,s}^*(\rho) \right]^2, \quad \forall s \in \mathcal{S} \quad (4.17)$$

where  $q_{rs}^*$  and  $f_{R,s}^*$  are the optimal solutions (as a function of  $\rho$ ) of problem (4.4, replacing  $\mathcal{G}$  with  $\bar{\mathcal{G}}$ ) and (4.10), respectively.

The relationship between maxinf point of  $W(\rho, \varphi)$  and the equilibrium spatial prices is stated in Lemma 4.4.3.

**Lemma 4.4.3.** (*equilibrium prices and maxinf-points*) For  $\rho \in \mathcal{S}$ ,  $\rho$  is a maxinf-point of the Walrasian function  $W$  such that  $W(\rho, \cdot) \geq 0$ , on  $\Delta_S$  if and only if  $\rho$  is an equilibrium point.

Proof. See B.1.

This conversion is inspired by a theoretical development on lopsided convergence of bifunction by Jofré and Wets (2009), which offers the flexibility of constructing a sequence of approximate bifunctions with desired properties (including convexity and continuity) to solve for the original maxinf-point of  $W(\rho, \varphi)$ . A similar approach has been successfully implemented in (Deride et al., 2019; Guo et al., 2016b). In this paper, we highlight the key steps of the algorithm design.

Since  $\inf_{\varphi} W(\rho, \varphi)$  lack of concavity in  $\rho$ , we apply an *augmented Lagrangian* for this non-concave formulation. Given sequences of nonnegative, nondecreasing scalars  $\{r^v\}$ ,  $\{M^v\}$ , one can define a sequence of *augmented Walrasian* functions for this problem as shown in Equation (4.18).

$$W^v(\rho, \varphi) = \inf_z \left\{ W(\rho, z) + \frac{1}{2r^v} |z - \varphi|^2 \mid z \in \Delta_S \right\}, \text{ on } [\rho^{v-1} - M^v, \rho^{v-1} + M^v] \times \Delta_S \quad (4.18)$$

The idea of this procedure is to approximate the problem of finding maxinf-points of the original Walrasian function  $W$ , by computation of approximate maxinf-points  $\rho_\varepsilon$  (see Definition 3) given

by a sequence of augmented Walrasians  $W^v$ , which are easier to solve. Note that the proposed approach is not a heuristic approach since the convergence of the proposed approximation scheme is guaranteed, as stated in Theorem 4.4.4.

**Definition 3.** (*approximating equilibrium point*)  $\rho_\varepsilon$  is an  $\varepsilon$ -approximating equilibrium point, denoted as  $\varepsilon - W$ , for  $\varepsilon \geq 0$ , if the following inequality holds:  $|W(\rho_\varepsilon, \cdot) - W| \leq \varepsilon$ .

**Theorem 4.4.4.** (*convergence of approximating maxinf-points*) Suppose that  $W$  is finite. Consider non-negative sequences  $\{r^v\}$ ,  $\{M^v\}$ , and  $\{\varepsilon^v\}$  such that  $r^v \rightarrow 0$ ,  $M^v \rightarrow 0$ ,  $\varepsilon^v \rightarrow 0$ . Let  $\{W^v\}$  be a family of augmented Walrasian functions associated with each parameters  $r^v$  and  $M^v$ . Let  $\rho^v \in \varepsilon^v - W^v$ , and  $\rho^*$  be any cluster point of  $\{\rho^v\}$ . Then  $\rho^* \in W$ .

Proof. See (Guo et al., 2016b).

Following Theorem 4.4.4, we propose the computational algorithm, as summarized in Algorithm 2, to achieve a sequence of approximating maxinf-points.

---

**Algorithm 2** Approximating Maxinf-Point Algorithm

---

$\rho^v$  initialization:  $v = 0$ ,  $\varepsilon^0$ ,  $M^0$ ,  $r^0$ ,  $\text{gap}^0$ ,  $\rho^0$ ,  $\varepsilon$ ,  $0 < c_1 < 1$ ,  $c_2 > 1$

**while**  $\text{gap}^v \geq \varepsilon$  **do**

Phase I: solve the minimization problem  $\varphi^{v+1} \in W^{v+1}(\rho^v, \cdot)$

Phase II: solve the maximization problem  $\rho^{v+1} \in W^{v+1}(\cdot, \varphi^{v+1})$

evaluate:  $\text{ES}_s(\rho^{v+1}), s \in \mathcal{S}$

let:  $\text{gap}^{v+1} = \max\{\text{ES}_s(\rho^{v+1}), s \in \mathcal{S}\}$

let:  $\varepsilon^{v+1} = c_1 \varepsilon^v$ ,  $M^{v+1} = c_2 M^v$ ,  $r^{v+1} = c_2 r^v$

let:  $v := v + 1$

---

Phase I consists of the minimization of a quadratic objective function over the  $S$ -dimensional simplex. This can be solved using Cplex/Gurobi solver. Phase II is done without considering first-order

information and relying on the BOBYQA algorithm (Powell, 2009), which performs a sequentially local quadratic fit of the objective functions, over box constraints, and solves it using a trust-region method. As  $r^V \rightarrow \infty$ , and  $M^V \rightarrow \infty$ , and  $\epsilon^V \rightarrow 0$ , in virtue of Theorem 4.4.4,  $\rho^V \rightarrow \rho^*$ , a maxinf-point of  $W$ .

**Remarks:**

1. The network augmenting procedures proposed in Section 4.3.2 result in a transportation network with a non-separable link-cost function (i.e.,  $t_D$  and  $t_R$ ), and the Jacobian of link-cost function is asymmetric, for which there is no known mathematical program whose solutions are the equilibrium flow pattern (Sheffi, 1985). But, the traffic equilibrium solutions can be achieved by the diagonalization method (Sheffi, 1985). We note that since the Jacobian of the link-cost functions is not positive definite when  $|a_1| \leq |a_2|$  and  $f_{D,s} \approx f_{R,s}$ , the convergence of traffic equilibrium is not guaranteed in theory, although the algorithm may still convergence when this condition is violated (Fisk and Nguyen, 1982). To make our algorithm robust to the non-convergence of traffic equilibrium, the algorithm (Algorithm 2) we develop here does not require convergence for every iteration.

2. Notice that problem (4.11) has at least  $|\mathcal{S}|$  trivial equilibrium solutions, where only one location  $s$  attracts all the drivers  $\sum_{r \in \mathcal{R}} Q_r$ . This is because when driver supply approaches 0 and price is finite at one location, rider demand will also approach 0 due to infinite waiting time, and vice versa. The insight from this is that the ride-sourcing market may need a critical mass of drivers and riders to sustain. In this paper, we only focus on non-trivial equilibrium.

**4.5 Numerical Examples**

In this section, we test our models and solution approaches on a three-node system (Figure 4.2) and Sioux Falls network (Figure 4.7). We implement our model on Pyomo 5.6.6 (Hart et al., 2017).

We solve the problems using IPOPT 3.12.13 (Wächter, 2009) for the “ignoring waiting time” case, and Cplex 12.8 & BOBYQA (in NLOpt 2.4.2) for the “considering waiting time ” case, both with 0.1% optimality gap. All the numerical experiments presented in this section were run on a 3.5 GHz Intel Core i5 processor with 8 GB of RAM memory, under Mac OS X operating system.

#### 4.5.1 Three-node Network

The three-node test instance (Figure 4.2) has one driver node (node 1) and two demand nodes (nodes 2 and 3). 50 drivers decide where to pick up riders at node 1.  $\beta_0 = 0, \beta_1 = 1, \beta_2 = 0.6$ . The link travel time has a form of  $t_a = t_a^0[1 + 0.15 * (v_a/c_a)^2]$ , where  $t_a^0$  is the free flow travel time and  $c_a$  is the link capacity parameter. The value of  $t_a^0$  and  $c_a$  for each link are shown in Figure 4.2. Notice that this test instance is symmetric except that links between nodes 1 and 3 have lower capacity than other links.

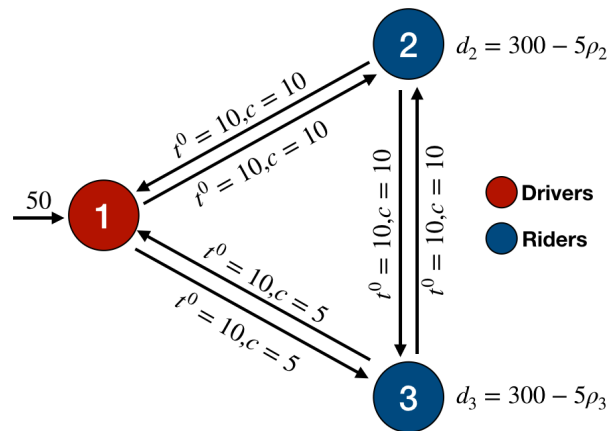


Figure 4.2: Three Nodes Test Network

#### 4.5.1.1 Ignoring Waiting Time

We assume both node 2 and 3 have a demand function of  $d_s = 300 - 5\rho_s$ . Solving this problem using our single-level reformulation (4.15), we have  $\rho_2 = 53.5, \rho_3 = 56.5$ . Notice that with these locational prices, the demand market at both nodes 2 and 3 are balanced. The difference between  $\rho_2$  and  $\rho_3$  is because of the link congestion between nodes 1 and 3. In other words, node 3 is more congested to travel to due to limited link capacity. This is because, in order to attract more drivers to pick up riders at node 3, TNC needs to offer a higher locational price. If TNC doesn't consider transportation congestion, the optimal locational prices are  $\rho'_2 = \rho'_3 = 55.0$ , which will lead to an unbalance of supply and demand on both demand nodes in reality. These optimal solutions are validated by applying global solver SCIP (Achterberg, 2009) directly on mixed-integer nonlinear programming (MINLP) reformulation of bi-level problem (4.5). Reformulating bi-level problem into MINLP is standard, so we omit the detailed procedures for conciseness. For the three-node example, our reformulation (4.15) can be solved by IPOPT in 0.1 seconds, while MINLP formulation takes SCIP 5.6 seconds to solve. Although both SCIP and IPOPT solve the problem for global optimal, MINLP formulation is sensitive to the selection of big-M parameters, which is a challenging issue for mixed integer programming with switching constraints (Guo et al., 2016a).

Because TNC aims to balance supply and demand, transportation congestion can be worse under spatial pricing. We compare two scenarios: surge pricing and uniform pricing. If we change from uniform pricing to surge pricing, the total travel time increases slightly from 1333 to 1336. The increase in total travel time is because TNC gives incentives (surge pricing) to drivers traveling through congested areas in order to balance demand. This result matches the empirical findings in (Nie, 2017) that ride-sourcing may increase transportation congestion compared to taxi services. We also investigated the sensitivity of different ride-sourcing penetration levels on the total system travel time, as shown in Figure 4.3. We can see that when the number of driver supply is at a

low level compared to total travel demand, the impact of increasing ride-sourcing market share on total travel time is less significant. However, with the increasing driver supply, the system's total travel time increases non-linearly. The main reason is the additional relocation and pick-up VMT added to the transportation system in addition to the original traveler VMT. Regulations may be needed to prevent pricing selfishly by TNCs to improve overall transportation mobility, especially in congested areas.

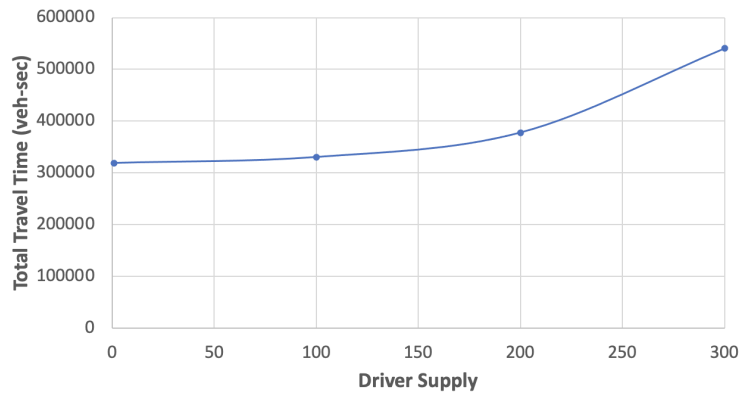


Figure 4.3: Impact of Ride-sourcing Service on Traffic Congestion

Next, we compare the impacts of TNC's objectives between maximizing profits and minimizing imbalance. The problem of maximizing profits is shown in (4.19).

$$\begin{aligned} &\text{maximize} && \sum_{s \in \mathcal{S}} \rho_s n_s && (4.19a) \\ &\rho \in \mathcal{S} \end{aligned}$$

subject to

$$n_s = \min\left(\sum_{r \in \mathcal{R}} q^{rs}, D^s - b^s \rho_s\right), \forall s \quad (4.19b)$$

$$(4.4) \quad (4.19c)$$

where:

$n_s$  : ride-sourcing matches at  $s$ .

(4.19a) maximizes total revenues. Since the majority of TNC operating costs go to marketing, administrative, and R&D <sup>5</sup>, which do not depend on the pricing decision variables. Therefore, the optimal pricing maximizes total revenues will also maximize total profits in the service territory. But we note that operating costs can be easily incorporated in (4.19a) if operating costs data is available. (4.19b) calculates the matches at each location  $s$ . For problem (4.19), we are not able to directly make the single-level reformulation as in (4.15). So we solve the bilevel problem (4.19) as MINLP using SCIP solver. The numerical results are shown in Figure 4.4a and Figure 4.4b, in which we conduct sensitivity analyses of the total number of drivers  $Q_r$  and interception of demand function  $D_s$  on equilibrium prices with different TNC objectives. When TNC aims to minimize imbalance for each location, the optimal prices are almost linearly decreasing (increasing) with increasing supply (demand) for this toy test instance. The trend of optimal prices is slightly more complex when TNC aims to maximize profits. When supply is relatively high or demand is relatively low, TNC's profits are not restricted by drivers' availability, so TNC will set prices converging to monopoly prices, which are  $D^s/2b^s, \forall s$ . This means that when a TNC aims to maximize profits, the supply and demand may not necessarily be balanced at each location. This finding is consistent with the results in (Bimpikis et al., 2019) where endogenously determined spatial prices aiming to maximize profits may not always balance supply and demand. Notice that when demand is comparable to supply, prices minimizing imbalance of ride requests are identical with prices maximizing TNC profits, which indicates that maximizing balance of supply and demand leads to maximum profits for TNC.

---

<sup>5</sup><https://news.crunchbase.com/news/understanding-uber-loses-money/>





Figure 4.4: Impacts of Supply and Demand on Optimal Surge Pricing

#### 4.5.1.2 Considering Waiting Time

When waiting time cannot be ignored for both drivers' and riders' decision making, we can solve problem (4.11) using Algorithm 2. We assume  $D_s = 300, \forall s$ . The riders' utility coefficients  $\beta'_0 = 0, \beta'_1 = 1, \beta'_2 = 0.6$ .

With different initial locational prices, the convergence pattern of prices and ES are shown in Figure 4.5. The average computation time is 27.8s, and all of the experiments converged efficiently to the same equilibrium solutions,  $\rho_2^* = 37.4, \rho_3^* = 38.0, f_{D,2}^* = f_{R,2}^* = 33.1, f_{D,3}^* = f_{R,3}^* = 16.9$ . Notice that TNC offers a slightly higher price at location 3, which is more congested to travel to. This observation is consistent with the observation in the "ignoring waiting time" case.

We compare the equilibrium solutions between the ignoring and considering waiting time cases,

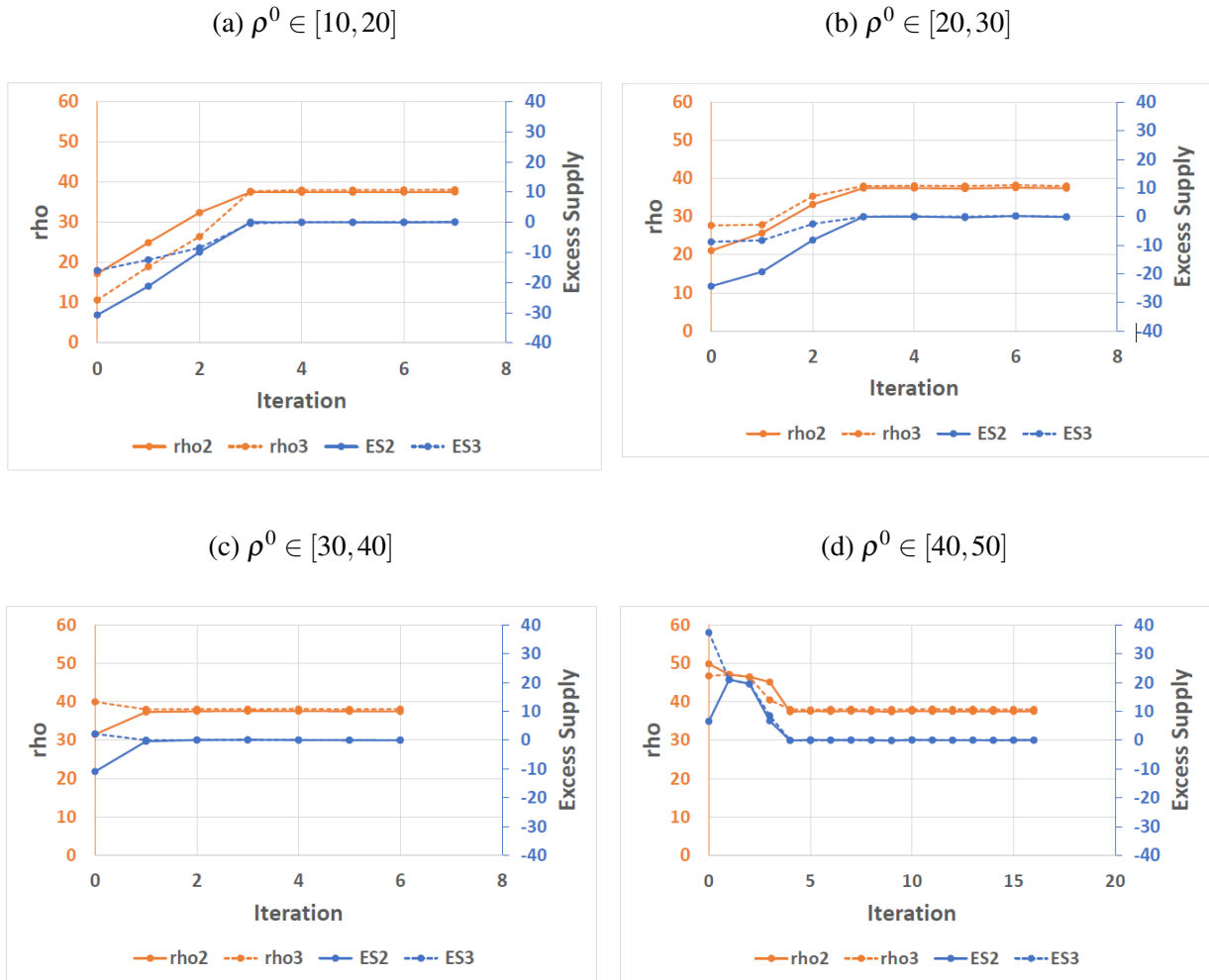


Figure 4.5: Convergence Patterns with Different Starting Points  $\rho^0$

with the results of equilibrium spatial pricing, driver supply, and rider demand shown in Figure 4.6. We can see that, compared to equilibrium prices, driver and rider flows are less sensitive to whether the waiting time is considered or not. The reason is that different locations have similar waiting times when supply and demand are balanced. For drivers, their relocation choices only depend on the utility difference between locations, which will not change if a similar waiting time is ignored. Therefore, driver supply remains unchanged (see Figure 4.6a). Given the similar amount of driver supply, TNC will need to raise up locational prices (see Figure 4.6c) to compensate for the ignored

waiting time, so that rider demand still balances driver supply (see Figure 4.6b).

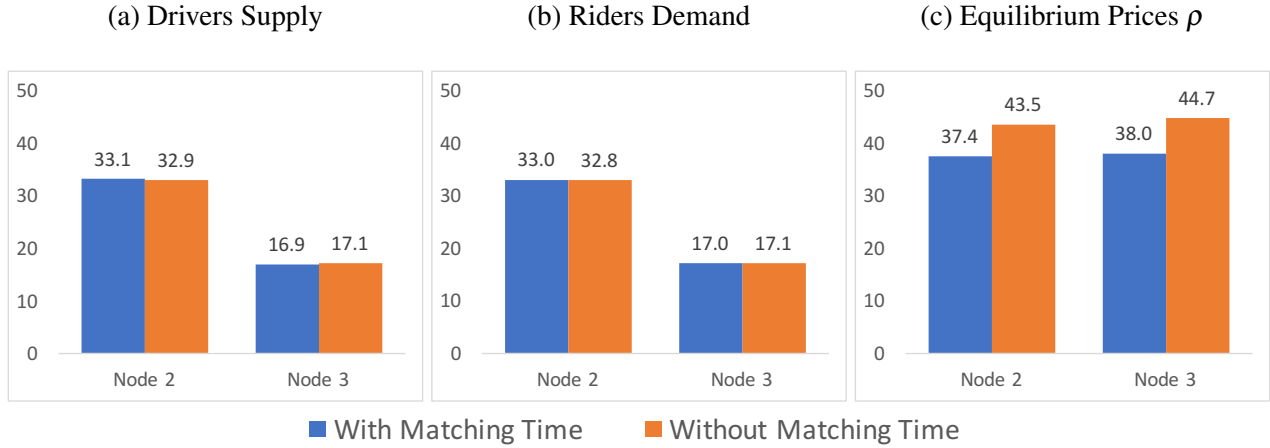


Figure 4.6: Equilibrium Solutions For Ignoring and Considering Waiting Time Cases

#### 4.5.2 Sioux Falls Network

We implement our model using a medium-size test system, Sioux Falls network<sup>6</sup>, as shown in Figure 4.7. In Figure 4.7, the red and blue nodes (12 of each) represent the original locations of drivers and riders, respectively. Drivers supply at each red node is 50, while the demand function at each blue node is  $d_s = 300 - 5\rho_s$ . We adopt a 4<sup>th</sup>-order Bureau of Public Roads (BPR) function for link cost:  $t_a = t_a^0[1 + 0.15 * (v_a/c_a)^4]$ . Coefficients of utility function for base case (4.1) are  $\beta_0 = 0, \beta_1 = 1, \beta_2 = 0.6$ .

##### 4.5.2.1 Ignoring Waiting Time

The optimal locational pricing decisions and total equilibrium travel time are shown in Figure 4.8. With price sensitivity coefficient  $\beta_2$  increase from 0.1 to 10, the price variances decrease (see

<sup>6</sup>Sioux Falls network consists of 24 nodes and 76 directed links. The number on each node/link in Figure 4.7 is the node/link index.

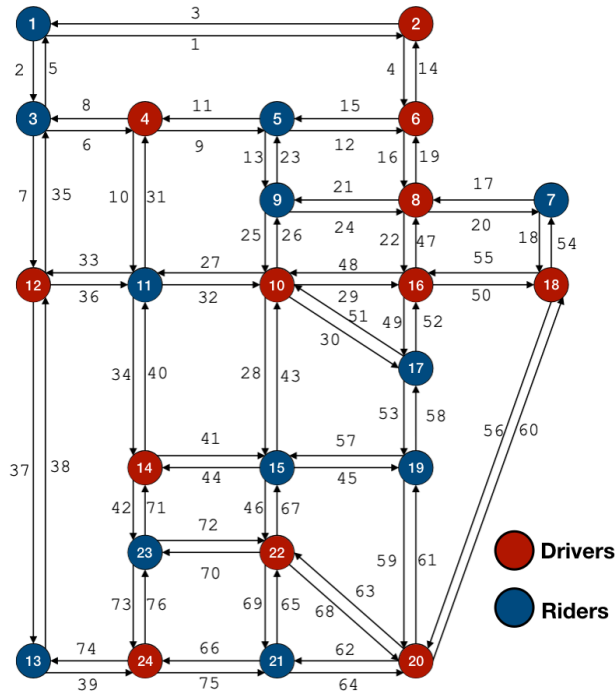


Figure 4.7: Sioux Falls Test Network

Figure 4.8a). This is because when drivers are more sensitive to prices, TNC needs a smaller price difference to attract drivers to the desired locations. At the same time, drivers are more willing to travel long distances to reach a location with higher prices. This will lead to higher system total travel time, as shown in Figure 4.8b.

SCIP is not able to solve the MINLP reformulation of problem (4.5) for Sioux Falls network, while IPOPT can efficiently solve (4.15) in 11.3, 6.1, and 4.9 seconds for  $b_2 = 0.1, 1, \text{ and } 10$ , respectively.

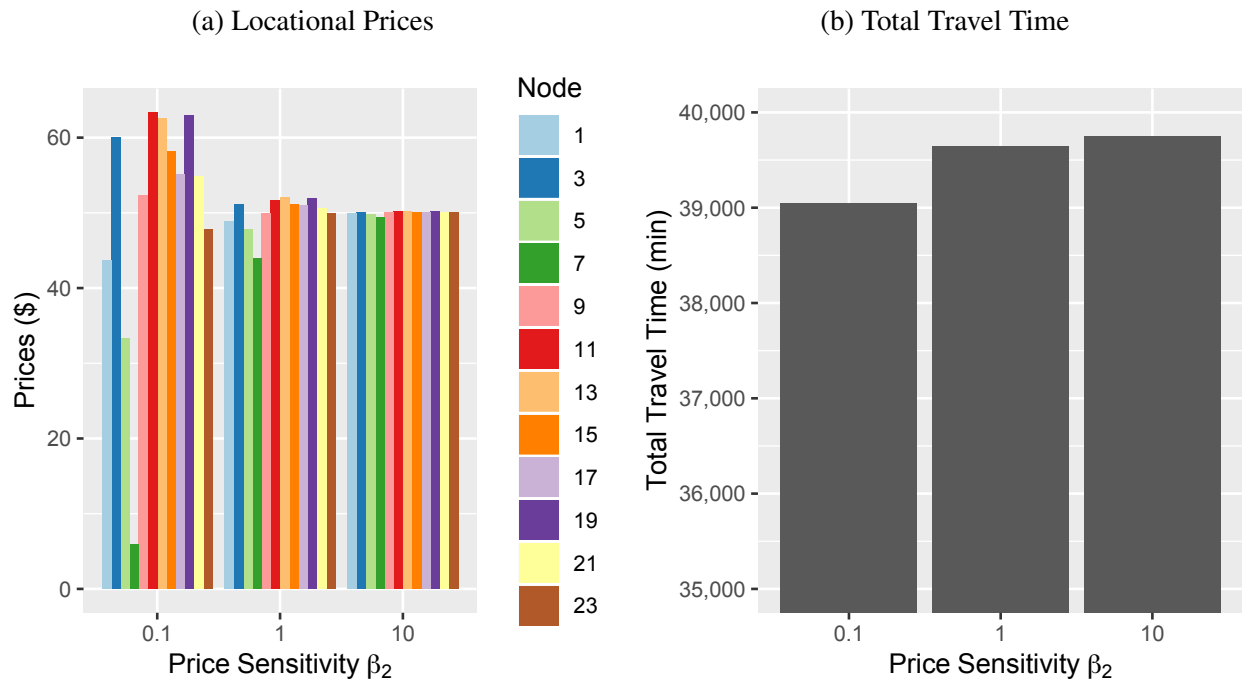


Figure 4.8: Impacts of Supply and Demand on Optimal Surge Pricing

#### 4.5.2.2 Considering Waiting Time

Considering the waiting time, the problem can be solved by Algorithm 2 in 6.4 hours, which is significantly longer than the computation time for the “ignoring waiting time” case. This is because of the lack of convexity for problem (4.11). But the algorithm converges reliably in 12 iterations, with convergence patterns of ES and  $\rho$  shown in Figure 4.9.

Comparing the impact of spatial pricing on transportation congestion, the total travel cost is 3805.5 when spatial pricing is adopted, which is higher than a total travel cost of 3770.4 when TNC adopts uniform pricing. This observation is again consistent with the results in the “ignoring waiting time” case.

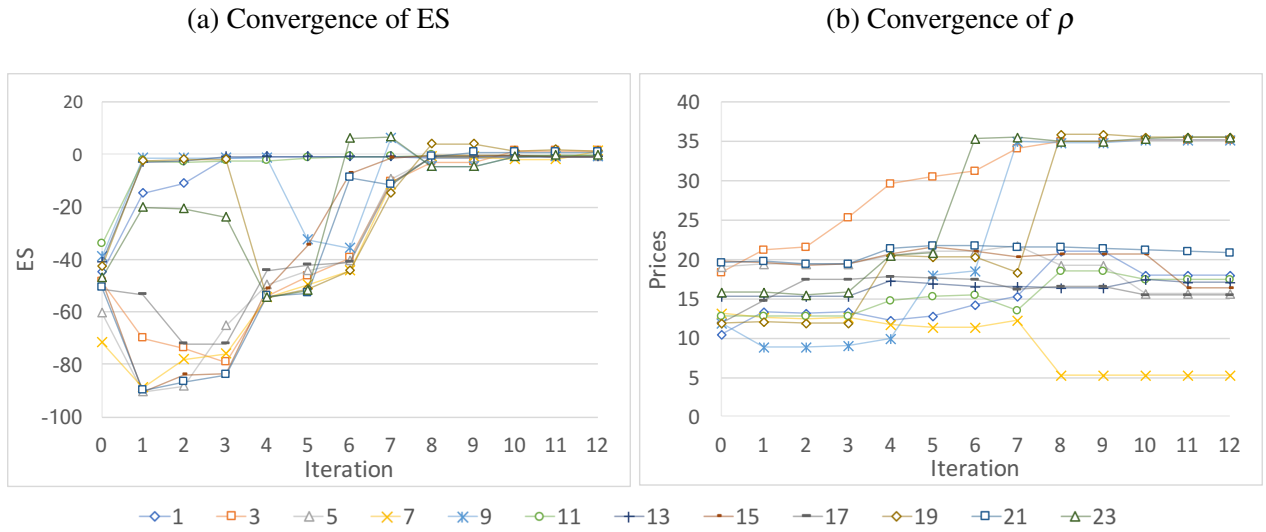


Figure 4.9: Convergence of ES and  $\rho$  Considering Waiting Time

#### 4.6 Discussion

In this paper, we have presented a new modeling framework and effective computational techniques for ride-sourcing spatial pricing problems, considering the leader-follower structure of TNCs, riders, and drivers in the peak traffic period. In the lower level, interactions between drivers' relocation, riders' mode choice, all travelers' routing behaviors, and network congestion are explicitly modeled. In the upper level, a single TNC provider optimally determines spatial prices to achieve its own objectives. When waiting time can be ignored, we show that the problem is equivalent to a convex optimization problem, which can be efficiently solved to global optimal by commercial nonlinear solvers. The existence and uniqueness of optimal solutions are proved. When waiting time cannot be ignored, we show that the lower-level drivers' and riders' two-sided interactions can be reformulated in a unified CDA framework with a non-separable link-cost function through network augmentation. The problem can be solved reliably by a lopsided convergence algorithm.

In addition to showing the effectiveness of our solution approaches, we demonstrated that (1) trans-

portation congestion affects spatial pricing strategies, and spatial pricing strategies may aggravate transportation congestion compared to uniform pricing; (2) spatial pricing strategies that minimize supply-demand imbalance also maximize TNC profits when maximum demand and supply are comparable; (3) drivers' sensitivity to surge multipliers will significantly impact spatial prices and transportation congestion; and (4) the impacts of waiting time may be small on the distribution of drivers and riders but may be large on the scale of spatial pricing.

This research can be extended in several directions. From a methodological viewpoint, our model mainly focuses on spatial aspects. Coupling temporal and spatial dimensions are critical to fully understand the impacts of dynamic pricing on transportation systems. A spatial-temporal model will be needed to fully capture other important characteristics of the ride-sourcing system, such as (1) partitioning of driver-hours (driving time with riders, repositioning time, idle waiting time, and pick-up time) ; (2) network effects of riders destination and travel time on the spatial and temporal distribution of available drivers. In addition, several assumptions in this study can be relaxed. For example, the assumption on single TNC and inelastic total driver supply could be relaxed to generalize our findings. Our augmented network approach may be further extended to consider multiple TNCs by duplicating the "market" nodes in Figure 4.1 to represent different service providers. However, single-level convex reformulation and equilibrium uniqueness may not be guaranteed. After gaining a clear understanding of the impacts of spatial pricing on transportation systems, strategies to mitigate the negative impacts from regulators' perspectives are urgently needed. Extending our model to a stochastic environment is another immediate next step considering the high level of uncertainties involved in the ride-sourcing market, such as congestion, rider demand, and driver supply. Although our model can be efficiently solved for small and medium transportation networks, algorithm development is needed for extremely large cases. Classic traffic assignment algorithms could be leveraged to solve the lower-level problem more efficiently. Solving a stochastic version of our model might also present further numerical challenges, for which decomposition

methods (such as scenario decomposition and Benders decomposition) developed for stochastic programs may be integrated with the reformulation techniques in this study.



## CHAPTER 5: CONCLUSIONS

Smart mobility, propelled by cutting-edge technologies, is heavily reliant on real-time information to enable informed decision-making, effective mode choice, dynamic vehicle control, optimized travel routing, and strategic vehicle relocation. Existing studies have demonstrated a general belief that increased information is advantageous for individual decision makers. However, it is vital to consider the contingent nature of the effects of information on transportation network performance, as some studies have indicated that while information sharing can be beneficial, there are instances where an excess of information may not necessarily enhance the safety and mobility of the entire system. This dissertation aims to examine the impact of information sharing among public and private transportation entities on system-level performance, addressing challenges associated with the absence of a comprehensive modeling framework that can accurately capture the decentralized multi-agent interaction in interconnected transportation networks while managing computational complexities arising from non-convexity and high dimensionality. To achieve this goal, this dissertation, in particular, presents novel modeling frameworks and computational solutions for three state-of-the-art smart mobility applications capable of sharing information with vehicles and investigates their impact on the transportation network. The first objective is to develop a scalable and novel model to assess the impact of en-route information sharing provided by infrastructure-to-vehicles (I2V) communication considering the adaptive capability of vehicles. The second objective is to develop a model to optimally control variable speed limits and assess its impact on network performance. The third objective is to develop a new modeling framework and computational technique to investigate the impact of information shared by private entities on transportation network congestion.

## 5.1 Summary of Major Results

- In the second chapter of this dissertation, we proposed a two-stage stochastic user equilibrium with recourse model that captures the adaptive routing behavior of connected and automated vehicles (CAVs) considering the availability of en-route information at various sharing locations provided by I2V communications. Through numerical experiments, we find that the optimal information sharing strategies depend on specific network configurations, and the effectiveness of information sharing is not solely determined by the quantity of information. Additionally, locational information sharing encourages traffic to pass through information nodes, enabling informed rerouting decisions. Furthermore, while mobility exhibits a monotonic decrease with increasing levels of origin-destination demand and incident severity, the relationship between network collision risks is more complex and can only be accurately assessed using a network modeling approach and real historical crash data.
- In the third chapter, we proposed a twin-delayed deep deterministic policy gradient model for the optimal control of variable speed limit controls (VSLCs) to optimize the network mobility and safety considering the rerouting behavior of vehicles. We utilized a four-node network to capture essential traffic dynamics in the presence of VSLC, which may not be readily observable in larger networks. The experimental results demonstrated that our proposed model outperforms various Deep Reinforcement Learning (DRL)-based models, resulting in substantial enhancements in network performance. Furthermore, we assessed the transferability of our proposed model by varying diverse traffic parameters and vehicle behaviors, showcasing its adaptability to changing traffic conditions. Our research also presented several traffic control strategies employing VSLC and underscored that conventional traffic engineering practices may negatively impact network mobility and safety when compared to controlling a routing link using VSLC. Lastly, we conducted a sensitivity analysis to determine the optimal placement of VSLC within the network for improving network

mobility and safety.

- In the fourth chapter, we proposed a novel modeling framework and efficient computational techniques for addressing spatial pricing problems in ride-sourcing services during peak traffic periods, considering the hierarchical structure of transportation network companies (TNCs), riders, and drivers, explicitly modeling interactions related to drivers' relocation, riders' mode choice, travelers' routing behaviors, and network congestion. Through numerical experiments, we demonstrated that transportation congestion impacts spatial pricing strategies and that such strategies can exacerbate congestion compared to uniform pricing. We also observed that spatial pricing strategies minimizing supply-demand imbalance can maximize TNC profits when demand and supply are comparable. The sensitivity of drivers to surge multipliers significantly influences spatial prices and transportation congestion. Furthermore, the impact of waiting time is found to have a small effect on the distribution of drivers and riders but can have a substantial impact on the scale of spatial pricing.

## **5.2 Future Research Direction**

During our investigation of the impact of information sharing on transportation networks, we recognized the broad and critical nature of this research topic, which presents numerous opportunities for further exploration. It is evident that there are additional research directions that warrant attention to deepen our understanding and contribute to the development of effective strategies and policies in the field of smart mobility. Some of the key areas for future investigation include:

- **Obtaining and dissemination of reliable Information:** In this dissertation, we examined the impact of information shared by various cutting-edge smart mobility applications. To delve deeper into the realm of smart mobility and its impact, it is important to explore several

intriguing avenues. One of the key aspects to consider is the acquisition of information. In smart mobility, there are various sources of information on transportation networks such as connected and autonomous vehicles, intelligent transportation systems, and the Internet of Things (IoT). Exploring the integration of information from these emerging technologies is crucial to understanding the impact of information. Data fusion techniques can play a crucial role in integrating diverse data sets, combining information from various sensors, sources, and modalities to provide a comprehensive and holistic view of the transportation network.

Following the acquisition of information, it is imperative to prioritize the security and integrity of data, placing particular emphasis on secure transmission and storage mechanisms. Thorough research on cybersecurity measures becomes essential to safeguard the confidentiality and integrity of collected information. This entails the implementation of encryption protocols, intrusion detection systems, and data integrity mechanisms as effective safeguards against potential cyber threats.

Simultaneously, privacy preservation emerges as a critical aspect to consider when dealing with sensitive information within the transportation domain. Exploring privacy-preserving mechanisms becomes essential to strike a balance between effective information sharing and protecting individuals' privacy. This could involve investigating techniques such as anonymization, differential privacy, and access control to uphold privacy while still enabling valuable insights from the collected data.

- **Understanding equity impact:** Another important aspect of understanding the impact of information is to assess the equity implications of information sharing in smart mobility systems. A thorough analysis should be conducted to understand how access to information influences various vehicle owners, including those with conventional vehicles and those with CAVs. The focus should be on evaluating potential disparities in information availability and their impact on transportation choices, service accessibility, and overall mobility out-

comes for different vehicle owners. For instance, information disparities can occur between owners of CAVs and conventional vehicles. CAV owners may have access to real-time traffic updates, route optimization, and predictive analytics tailored specifically to their vehicle type. In contrast, conventional vehicle owners might not have access to such advanced information, leading to potential disparities in route choices, travel times, and overall driving experience. By exploring the design of information-sharing mechanisms, the aim is to promote equity among vehicle owners and mitigate any potential inequalities. The objective is to create an information-sharing ecosystem that fosters fairness and equal opportunities for all vehicle owners, regardless of the type of vehicle they own.

- **Multi-modal information sharing:** Investigating the impact of sharing information across different modes of transportation, considering the interactions and interdependencies among different transportation systems is an important research direction in the field of smart mobility. In future research on the impact of multi-modal information sharing in smart mobility, a primary focus should be placed on understanding the dynamics and interactions among modes, and optimization of multi-modal information sharing, including route planning, mode choice, and transfer coordination, to improve overall system performance and provide seamless travel experiences.

The impact of information sharing in smart mobility is a broad and critical research topic with immense potential for future investigation. This dissertation on the impact of information sharing on transportation networks has laid a solid foundation for comprehending the significance of this phenomenon. By investigating the effects of information sharing in the context of smart mobility, our research has contributed valuable insights into the transformative potential of this practice. By building upon the findings and methodologies presented in this dissertation, future studies can delve deeper into the intricacies of information sharing and its implications for transportation systems.

**APPENDIX A: PROOF OF SYSTEM-LEVEL IMPACTS OF EN-ROUTE  
INFORMATION SHARING CONSIDERING ADAPTIVE ROUTING**

### A.1 Proof of Theorem 2.3.1

Firstly, the objective function (2.3a) is convex because  $t_a$  is monotone increasing with respect to traffic flow. In addition, the constraints for problem (2.3) are all linear. Therefore, the optimization problem (2.3) is convex. The Lagrangian function of problem (2.3) is as follows:

$$\begin{aligned} \mathcal{L}(x_{a,k}^{rs}, x_p(\xi), \lambda_{a,k}^{rs}(\xi), \gamma^{rs}(\xi)) &= \mathbb{E}_\xi \sum_{a \in \mathcal{A}} \int_0^{v_a(\xi)} t_a(u, \xi) du \\ &\quad - \sum_{rs \in RS} \sum_{\xi \in \Xi} \gamma^{rs}(\xi) \left[ \sum_{p \in \mathcal{P}^{rs}} x_p(\xi) - q^{rs} \right] \\ &\quad - \sum_{rs \in RS} \sum_{a \in \mathcal{A}} \sum_{\xi \in \Xi} \lambda_{a,k}^{rs}(\xi) \left[ \sum_{p \in \mathcal{P}_k^{rs}} \delta_{ap}^+ x_p(\xi) - x_{a,k}^{rs} \right] \end{aligned}$$

The first-order derivatives of  $\mathcal{L}$  with respect to  $x_{a,k}^{rs}$  and  $x_p(\xi)$  are as follows:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_{a,k}^{rs}} &= \sum_{\xi \in \Xi} \lambda_{a,k}^{rs}(\xi), \quad \forall a, k, rs \\ \frac{\partial \mathcal{L}}{\partial x_p(\xi)} &= Pr_\xi \sum_{a \in \mathcal{A}} \delta_{ap} t_a(v_a(\xi), \xi) - \sum_{a \in \mathcal{A}} \delta_{ap}^+ \lambda_{a,k}^{rs}(\xi) - \gamma^{rs}(\xi), \quad \forall p, \xi \end{aligned}$$

where  $Pr_\xi$  is the probability measurement of scenario  $\xi$ .

Based on the Karush–Kuhn–Tucker (KKT) theorem, the optimality conditions of problem (2.3) is equivalent to the following complementarity conditions in additions to constraints (2.3b ~ 2.3c):

$$0 \leq x_{a,k}^{rs} \quad \perp \quad \sum_{\xi \in \Xi} \lambda_{a,k}^{rs}(\xi) \geq 0, \quad \forall a, k, rs \quad (\text{A.1a})$$

$$0 \leq x_p(\xi) \quad \perp \quad Pr_\xi \sum_{a \in \mathcal{A}} \delta_{ap} t_a(v_a(\xi), \xi) - \sum_{a \in \mathcal{A}} \delta_{ap}^+ \lambda_{a,k}^{rs}(\xi) - \gamma^{rs}(\xi) \geq 0, \quad \forall p, \xi \quad (\text{A.1b})$$

### A.1.1 Sufficient Condition of Theorem 2.3.1

**We first prove the sufficient condition by showing that every optimal solution of optimization problem (2.3) is a Stochastic User Equilibrium through the following two steps.**

Step 1: We first show that all the used subpaths in the second stage have the same cost, which is no more than the travel cost of the unused subpaths.

For any given path  $p$ , we separate the links before and after the first information node into two groups:  $\mathcal{A}_p^+$  and  $\mathcal{A}_p^-$ . Notice that for the paths do not pass through any information node,  $\mathcal{A}_p^-$  will be empty set and  $\mathcal{A}_p^+$  will contain all the links of path  $p$ . Furthermore, for those paths  $p$  that share the same set of  $\mathcal{A}_p^+$  and same  $rs$ , we group them into hyperpath set  $P_k^{rs}$ , with  $k$  denoting the  $k^{\text{th}}$  hyperpath set for  $rs$ .

For a given 3-tuple  $(k, rs, \xi)$  and  $\forall p \in P_k^{rs}$ , based on (A.1b) we have:

$$\{Pr_\xi [ \sum_{a \in \mathcal{A}_p^+} t_a(v_a(\xi), \xi) + \sum_{a \in \mathcal{A}_p^-} t_a(v_a(\xi), \xi) ] - \sum_{a \in \mathcal{A}_p^+} \lambda_{a,k}^{rs}(\xi) - \gamma^{rs}(\xi) \} (x_p(\xi) - x_p^*(\xi)) \geq 0 \quad (\text{A.2})$$

Because we consider  $p \in P_k^{rs}$ , by definition, these paths have the same  $\mathcal{A}_p^+$ . We denote  $Pr_\xi \sum_{a \in \mathcal{A}_p^+} t_a(v_a(\xi), \xi) - \sum_{a \in \mathcal{A}_p^+} \lambda_{a,k}^{rs}(\xi) \doteq t_{k,rs}^+(\xi)$ . On one hand, for any  $p \in P_k^{rs}$  that  $x_p(\xi) > 0$ , (A.2) can be simplified into:

$$\sum_{a \in \mathcal{A}_p^-} t_a(v_a(\xi), \xi) = \frac{1}{Pr_\xi} (\gamma^{rs}(\xi) - t_{k,rs}^+(\xi)) \quad (\text{A.3})$$

where the right hand side is independent of  $p$ . On the other hand, for those  $p \in P_k^{rs}$  that  $x_p(\xi) = 0$ , (A.2) can be simplified into:



$$\sum_{a \in \mathcal{A}_p^-} t_a(v_a(\xi), \xi) \geq \frac{1}{Pr_\xi} (\gamma^{rs}(\xi) - t_{k,rs}^+(\xi)) \quad (\text{A.4})$$

Combining (A.3) and (A.4), we have shown that all the used subpaths in the second stage have the same cost, which is no more than the travel cost for the unused subpaths.

Step 2: Then we show that the expected travel cost for all the first stage used paths connecting a pair of origin  $r$  and destination  $s$  are the same, which is no more than the travel cost for the unused paths connecting  $rs$ .

For a given 2-tuple  $(k, rs)$ , if there exist  $a \in \mathcal{A}$  such that  $x_{a,k}^{rs} > 0$ , we have  $\sum_{\xi \in \Xi} \lambda_{a,k}^{rs}(\xi) = 0$  based on (A.1a). Because of (2.3d), for any scenario  $\xi \in \Xi$ , there exists at least one path, denoted as  $p(\xi) \in P_k^{rs}$ , such that  $x_{p(\xi)}(\xi) > 0$ . Therefore, based on (A.1b), we have the following:

$$Pr_\xi \sum_{a \in \mathcal{A}} \delta_{ap(\xi)} t_a(v_a(\xi), \xi) - \sum_{a \in \mathcal{A}} \delta_{ap(\xi)}^+ \lambda_{a,k}^{rs}(\xi) - \gamma^{rs}(\xi) = 0 \quad (\text{A.5})$$

Take summation of (A.5) over all  $\xi \in \Xi$ , we have:

$$\sum_{\xi \in \Xi} [Pr_\xi \sum_{a \in \mathcal{A}} \delta_{ap(\xi)} t_a(v_a(\xi), \xi) - \sum_{a \in \mathcal{A}} \delta_{ap(\xi)}^+ \lambda_{a,k}^{rs}(\xi) - \gamma^{rs}(\xi)] = 0 \quad (\text{A.6})$$

Because of the following two facts:

1.  $\sum_{\xi \in \Xi} \lambda_{a,k}^{rs}(\xi) = 0$  if  $x_{a,k}^{rs} > 0$ ;
2.  $\delta_{ap(\xi)}^+$  is identical for all  $\xi$  as long as  $p(\xi) \in P_k^{rs}$ ;

we have  $\sum_{\xi \in \Xi} \sum_{a \in \mathcal{A}} \delta_{ap(\xi)}^+ \lambda_{a,k}^{rs}(\xi) = \sum_{a \in \mathcal{A}} \delta_{ap(\xi)}^+ \sum_{\xi \in \Xi} \lambda_{a,k}^{rs}(\xi) = 0$ . Therefore, (A.6) becomes:

$$\sum_{\xi \in \Xi} Pr_\xi \sum_{a \in \mathcal{A}} \delta_{ap(\xi)} t_a(v_a(\xi), \xi) = \sum_{\xi \in \Xi} \gamma^{rs}(\xi) \quad (\text{A.7})$$

Notice that the left hand side of (A.7) represent the expected travel cost of first stage used path  $k$  and the right hand side of (A.7) is independent of  $k$ , which means the expected travel cost of all the first stage used paths connecting  $rs$  are identical.

For a given 2-tuple  $(k, rs)$ , if  $\forall a \in \mathcal{A}$ ,  $x_{a,k}^{rs} = 0$ , which means  $k^{th}$  first stage path is unused for connecting  $rs$ , we have  $\sum_{\xi \in \Xi} \lambda_{a,k}^{rs}(\xi) \geq 0$ . Same processes as above, we will have:

$$\sum_{\xi \in \Xi} Pr_{\xi} \sum_{a \in \mathcal{A}} \delta_{ap(\xi)} t_a(v_a(\xi), \xi) = \sum_{a \in \mathcal{A}} \delta_{ap(\xi)}^+ \sum_{\xi \in \Xi} \lambda_{a,k}^{rs}(\xi) + \sum_{\xi \in \Xi} \gamma^{rs}(\xi) \geq \sum_{\xi \in \Xi} \gamma^{rs}(\xi) \quad (\text{A.8})$$

Till now, we have finished our proof on the sufficient condition of Theorem 2.3.1.

#### A.1.2 Necessary Condition of Theorem 2.3.1

**Secondly, we prove the necessary condition by showing that every Stochastic User Equilibrium is one of the optimal solutions of optimization problem (3).**

Based on Definition 1, we have the following two variational inequalities:

$$\left\{ \sum_{\xi \in \Xi} Pr_{\xi} \left[ \sum_{a' \in \mathcal{A}_p^+} t_{a'}(v_{a'}^*(\xi), \xi) + t_k^-(\xi) \right] - t^{rs} \right\} (x_{a,k}^{rs} - x_{a,k}^{rs*}) \geq 0, \forall a \in \mathcal{A}_p^+, p \in P_k^{rs}, k \in K^{rs}, rs \quad (\text{A.9})$$

$$\left[ \sum_{a \in \mathcal{A}_p^-} t_a(v_a^*(\xi), \xi) - t_k^-(\xi) \right] (x_p(\xi) - x_p^*(\xi)) \geq 0, \forall p \in P_k^{rs}, k \in K^{rs}, rs, \xi \quad (\text{A.10})$$

where  $t_k^-(\xi)$  is the travel cost after receiving information and  $t^{rs}$  is the expected travel cost.

From VI (A.10), we have the following:

$$\left[ \sum_{a \in \mathcal{A}_p} t_a(v_a^*(\xi), \xi) - \sum_{a \in \mathcal{A}_p^+} t_a(v_a^*(\xi), \xi) - t_k^-(\xi) \right] (x_p(\xi) - x_p^*(\xi)) \geq 0, \forall p \in P_k^{rs}, k \in K^{rs}, rs, \xi \quad (\text{A.11})$$

Construct  $\lambda_{a,k}^{rs}(\xi)$  and  $\gamma^{rs}(\xi)$  in the following way:

$$\gamma^{rs}(\xi) \doteq Pr_\xi t^{rs} \quad (\text{A.12})$$

$$\lambda_{a,k}^{rs}(\xi) \doteq \frac{1}{|\mathcal{A}_p^+|} \{ Pr_\xi [ \sum_{a \in \mathcal{A}_p^+} t_a(v_a^*(\xi), \xi) + t_k^-(\xi) ] - \gamma^{rs}(\xi) \} \quad (\text{A.13})$$

so that:

$$\sum_{\xi \in \Xi} \gamma^{rs}(\xi) = t^{rs} \quad (\text{Because } \sum_{\xi} Pr_\xi = 1) \quad (\text{A.14})$$

$$\sum_{a \in \mathcal{A}_p^+} \lambda_{a,k}^{rs}(\xi) + \gamma^{rs}(\xi) \doteq Pr_\xi [ \sum_{a \in \mathcal{A}_p^+} t_a(v_a^*(\xi), \xi) + t_k^-(\xi) ] \quad (\text{A.15})$$

Plug equation (A.15) in VI (A.11), we will have condition (A.1b).

Plug equation (A.15) in VI (A.9), we will have:

$$\left\{ \sum_{\xi \in \Xi} [ \sum_{a' \in \mathcal{A}_p^+} \lambda_{a',k}^{rs}(\xi) + \gamma^{rs}(\xi) ] - t^{rs} \right\} (x_{a,k}^{rs} - x_{a,k}^{rs*}) \geq 0, \forall a \in \mathcal{A}_p^+, p \in P_k^{rs}, k \in K^{rs}, rs \quad (\text{A.16})$$

Because of (A.14), VI (A.16) can be simplified to:

$$\left\{ \sum_{\xi \in \Xi} \sum_{a' \in \mathcal{A}_p^+} \lambda_{a',k}^{rs}(\xi) \right\} (x_{a,k}^{rs} - x_{a,k}^{rs*}) \geq 0, \forall a \in \mathcal{A}_p^+, p \in P_k^{rs}, k \in K^{rs}, rs \quad (\text{A.17})$$

Based on our construction (A.13),  $\lambda_{a',k}^{rs}$  is independent of  $a'$ . So VI (A.17) is equivalent to condition (A.1a).

Till now, we have finished our proof on the necessary condition of Theorem 2.3.1.

**APPENDIX B: PROOF , EXPLANATION AND EXTENSION OF  
SPATIAL PRICING OF RIDE-SOURCING SERVICES IN CONGESTED  
TRANSPORTATION NETWORK**

## B.1 Proofs.

Proof (Lemma 4.3.1). The proof of Lemma 4.3.1 is outlined as follows. Since (4.4a) is strictly convex (assuming  $t_a(\cdot)$  is a strictly monotone increasing function) and constraints of model (4.4) are linear, Karush-Kuhn-Tucker (KKT) conditions are the sufficient and necessary optimality conditions, as shown in (B.1a) - (B.1c) and (4.4b) & - (4.4e), where  $a_i$  and  $a_j$  are starting and ending nodes of link  $a$ .

$$0 \leq q_{rs} \perp \frac{1}{\beta_1} (\ln q_{rs} - \beta_2 \rho_s - \beta_{0,s}) + (\hat{\lambda}_{rs}^r - \hat{\lambda}_{rs}^s) - \gamma_r \geq 0 \quad (\text{B.1a})$$

$$0 \leq \hat{v}_{rs}^a \perp t_a(v_a) - \hat{\lambda}_{rs}^{a_i} + \hat{\lambda}_{rs}^{a_j} \geq 0 \quad (\text{B.1b})$$

$$0 \leq \check{v}_{rs}^a \perp t_a(v_a) - \check{\lambda}_{sk}^{a_i} + \check{\lambda}_{sk}^{a_j} \geq 0 \quad (\text{B.1c})$$

$$(4.4b) - (4.4e)$$

We need to show that the KKT conditions (B.1) are equivalent to (i) the Wardrop equilibrium conditions (i.e., travel time in all routes actually used are equal and less than those that would be experienced by a single vehicle on any unused route) and (ii) logit model solutions. We outline the proofs of (i) and (ii) as follows.

Proof of (i): On one hand, for any path  $p$  connecting  $r$  and  $s$  with positive flow (i.e.  $x_p > 0$ ),  $\forall a \in p$ ,  $\hat{v}_{rs}^a > 0$ . Because of (B.1b),  $t_a(v_a) = \hat{\lambda}_{rs}^{a_i} - \hat{\lambda}_{rs}^{a_j}$ . Path travel time  $\sum_{a \in p} t_a(v_a) = \hat{\lambda}_{rs}^r - \hat{\lambda}_{rs}^s$ . On the other hand, for any path  $p'$  connecting  $r$  and  $s$  with zero flow (i.e.  $x_{p'} = 0$ ),  $\sum_{a \in p'} t_a(v_a) \geq \hat{\lambda}_{rs}^r - \hat{\lambda}_{rs}^s$ . Similarly, we can show the Wardrop equilibrium conditions for OD pairs  $s$  and  $k$  using (B.1c). Therefore, the traffic flows achieved by optimization problem (4.4) satisfy the Wardrop equilibrium conditions.

Proof of (ii): based on (B.1a), note that  $q_{rs} > 0$ ,  $\frac{1}{\beta_1}(\ln q_{rs} - \beta_2 \rho_s - \beta_{0,s}) + (\hat{\lambda}_{rs}^r - \hat{\lambda}_{rs}^s) - \gamma_r = 0$ , which means  $q_{rs} = e^{\beta_{0,s} - \beta_1(\hat{\lambda}_{rs}^r - \hat{\lambda}_{rs}^s) + \beta_2 \rho_s + \gamma_r} = e^{U_{rs} + \gamma_r}$ . Because of (4.4e),  $Q_r = \sum_{s \in \mathcal{S}} q_{rs} = e^{\gamma_r} \sum_{s \in \mathcal{S}} e^{U_{rs}}$ . So  $e^{\gamma_r} = \frac{1}{\sum_{s \in \mathcal{S}} e^{U_{rs}}} Q_r$ . We have  $q_{rs} = \frac{e^{U_{rs}}}{\sum_{s \in \mathcal{S}} e^{U_{rs}}} Q_r$ , which follows the Logit model results.

Proof (Lemma 4.4.1). Denote the optimized drivers distribution of model (4.4) as  $q_{rs}^*$ . Because of constraint (4.4e), the total driver supply is:

$$\sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{S}} q_{rs}^* = \sum_{r \in \mathcal{R}} Q_r \doteq \bar{Q},$$

which is a constant.

Because of constraint (4.5b),  $m_s \leq \sum_{r \in \mathcal{R}} q_{rs}^*$ . Therefore,  $\sum_{s \in \mathcal{S}} m_s \leq \sum_{s \in \mathcal{S}} \sum_{r \in \mathcal{R}} q_{rs}^* = \bar{Q}$ . Denote the optimal value of  $m_s$  as  $m_s^*$ , we have:

$$\sum_{s \in \mathcal{S}} m_s^* \leq \bar{Q} \tag{B.2}$$

On the other hand,

$$\sum_{s \in \mathcal{S}} m_s^* \geq 0 \tag{B.3}$$

The equality in (B.3) holds if and only if

$$\sum_{r \in \mathcal{R}} q_{rs}^* = D_s - b_s \rho_s, \forall s \in \mathcal{S} \tag{B.4}$$

In other words, we just need to show that  $\exists \rho \in \mathcal{S}$ , s.t. (B.4) holds.

Define excess supply  $ES^s(\rho) = \sum_{r \in \mathcal{R}} q_{rs}^*(\rho) - (D_s - b_s \rho_s), \forall s$ . Denote  $\underline{\rho}^s = (D^s - \bar{Q})/b^s$  and  $\bar{\rho}_s = D_s/b_s$ . Define a convex compact set  $H = \Pi_{s \in \mathcal{S}} [\underline{\rho}_s, \bar{\rho}_s]$ . We construct a function  $\rho' = h(\rho) =$

$(h_s(\rho_s), \forall s)$ , such that:

$$\rho'_s = h_s(\rho_s) = \rho_s - \frac{(\bar{\rho}_s - \underline{\rho}_s)ES_s(\rho)}{\bar{Q}} = \rho_s - \frac{ES_s(\rho)}{b_s}, \rho_s \in [\underline{\rho}_s, \bar{\rho}_s]$$

We would like to show that  $\rho'_s \in [\underline{\rho}_s, \bar{\rho}_s]$ . Firstly, it can be easily seen that  $h_k(\underline{\rho}_s) = \underline{\rho}_s - \frac{ES_s(\underline{\rho})}{b_s} = \underline{\rho}_s - \frac{\sum_{r \in \mathcal{R}} q_{rs}^*(\underline{\rho}) - (D_s - b_s \underline{\rho}_s)}{b_s} \leq \frac{D_s - \bar{Q}}{b_s} - \frac{0 - \bar{Q}}{b_s} = \bar{\rho}_s$ , and  $h_k(\bar{\rho}_s) = \bar{\rho}_s - \frac{ES_s(\bar{\rho})}{b_s} = \bar{\rho}_s - \frac{\sum_{r \in \mathcal{R}} q_{rs}^*(\bar{\rho}) - (D_s - b_s \bar{\rho}_s)}{b_s} \geq \frac{D_s}{b_s} - \frac{(Q-0)}{b_s} = \underline{\rho}_s$ . Secondly, we only need to show  $h_s(\rho_s)$  is a monotone decreasing function. To show that, we take first order derivative of  $h_s(\rho_s)$  with respect to  $\rho_s$ :

$$\dot{h}_s(\rho_s) = 1 - \frac{1}{b_s} \frac{\partial ES_s(\rho)}{\partial \rho_s} = 1 - \frac{1}{b_s} \left( \frac{\partial \sum_{r \in \mathcal{R}} q_{rs}^*(\rho)}{\partial \rho_s} - \frac{\partial (D_s - b_s \rho_s)}{\partial \rho_s} \right)$$

Because  $\frac{\partial \sum_{r \in \mathcal{R}} q_{rs}^*(\rho)}{\partial \rho_s} \geq 0$ ,

$$\dot{h}_s(\rho_s) \leq 1 - \frac{1}{b_s} (0 + b_s) = 0$$

Therefore, we have proved function  $h(\cdot)$  is a mapping from  $H$  to  $H$ . In addition, because  $h(\cdot)$  is monotone and continuous; and  $H$  is a convex compact subset of Euclidean space, based on Brouwer's fixed-point theorem, there exists  $\rho^* \in H$ , such that  $\rho^* = h(\rho^*)$ . Therefore, for any  $s$ ,  $\exists \rho_s^*$  such that  $\rho_s^* = \rho_s^* - \frac{ES_s(\rho^*)}{b_s}$ , which leads to  $ES_s(\rho^*) = 0, \forall s$ .

To show the uniqueness of  $\rho^*$ , assuming that there exist  $\rho^{**}$  such that  $ES_s(\rho^{**}) = 0, \forall s$ . Therefore,  $h(\rho^{**}) = \rho^{**}$ , i.e.,  $\rho^{**}$  is another fixed point of  $\rho' = h(\rho)$ . Without loss of generality, assuming  $\rho_s^{**} > \rho_s^*$ . Since  $h_s(\cdot)$  is monotone decreasing,  $\rho_s^{**} = h_s(\rho_s^{**}) \leq h_s(\rho_s^*) = \rho_s^*$ , which is a contraction with the assumption that  $\rho_s^{**} > \rho_s^*$ .

Proof (Theorem 4.4.2).

Because problem (4.15), (4.4) and (4.13) are convex optimization, with convex and continuously



differentiable objective function, closed and convex constraint sets, the optimal solutions to (4.15), (4.4) and (4.13) are equivalent to the solutions to certain variational inequalities (Nagurney and Siokos, 1997).

Denote the following items for notation conciseness.

- $\mathcal{X}$ : constraint set (4.4b - 4.4e);
- $\{x, q\}$ : decision variables of model (4.4), where  $x$  include all the decision variables except  $q$ ;
- $f(x, q)$ : objective function of model (4.4) excluding  $\sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{S}} q_{rs} \rho_s$ . That is  $f(x, q) = \frac{\beta_1}{\beta_2} \sum_{a \in \mathcal{A}} \int_0^{v_a} t_a(v_a) du + \frac{1}{\beta_2} \sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{S}} q_{rs} (\ln q_{rs} - 1 - \beta_{0,s})$ ;
- $\{d\}$ : decision variables of model (4.13);
- $g(d)$ : objective function of model (4.13) excluding  $\sum_{s \in \mathcal{S}} \rho_s d_s$ . That is  $g(d) = \sum_{s \in \mathcal{S}} 1/b_s [-D_s d_s + \frac{1}{2}(d_s)^2]$ ;

Following these notation, model (4.15) can be denoted as:

$$\min_{x, q, d \geq 0} f(x, q) + g(d), \text{ s.t. } \{x, q\} \in \mathcal{X}, \sum_{r \in \mathcal{R}} q_{rs} = d_s, \forall s$$

The equivalent variational inequalities for model (4.4), (4.13), (4.14) are shown in (B.5), (B.6), and (B.7), respectively.

$$\nabla_x f(x^*, q^*)(x - x^*) + \nabla_q f(x^*, q^*)(q - q^*) - \sum_{r, s} \rho_s^* (q_{rs} - q_{rs}^*) \geq 0, \forall x, q \in \mathcal{X}, x, q \geq 0 \quad (\text{B.5})$$

$$\nabla_d g(d^*)(d - d^*) + \sum_s \rho_s^* (d_s - d_s^*) \geq 0, \forall d \geq 0 \quad (\text{B.6})$$

$$\left( \sum_{r \in \mathcal{R}} q_{rs}^* - d_s^* \right) (\rho - \rho^*) \geq 0, \forall \rho \quad (\text{B.7})$$

On the other hand, model (4.15) can be solved by relaxing constraint (4.15b), whose KKT conditions are equivalent to VI (B.8).

$$\begin{aligned}
& \nabla_x f(x^*, q^*)(x - x^*) + \nabla_q f(x^*, q^*)(q - q^*) + \nabla_d g(d^*)(d - d^*) \\
& - \sum_{r,s} \rho_s^* (q_{rs} - q_{rs}^*) + \sum_s \rho_s^* (d_s - d_s^*) + \left( \sum_{r \in \mathcal{R}} q_{rs}^* - d_s^* \right) (\rho - \rho^*) \geq 0 \\
& \forall x, q \in \mathcal{X}, x, q, d \geq 0
\end{aligned} \tag{B.8}$$

Next, we just need to show that combining (B.5, B.6, B.7) is equivalent to (B.8). To show that, we will explain for any solutions satisfy (B.5, B.6, B.7), they also satisfy (B.8), and vice versa.

For solutions  $(x^*, q^*, d^*, \rho^*)$  satisfying (B.5, B.6, B.7), they also satisfy (B.8) by combining VI (B.5, B.6, B.7). On the other hand, for  $(x^*, q^*, d^*, \rho^*)$  satisfying (B.8), by selecting  $d = d^*, \rho = \rho^*$ , (B.8) becomes (B.5). Similarly, (B.6) and (B.7) are also satisfied.

Note that we do not derive VIs for the bilevel problem (4.5) when we prove Theorem 4.4.2. The reason why bilevel formulation (4.5) is equivalent to solving (4.4), (4.13), and (4.14) simultaneously is because of Lemma 4.4.1. In Lemma 4.4.1, we showed that the optimal pricing solutions to bilevel problem (4.5) are equivalent to market clearing prices for a single-level multi-agent optimization problem with (market) equilibrium constraints (MOPEC). The MOPEC problem considered here has two dummy decision makers: "drivers" (collectively) solve model (4.4), "riders" (collectively) solve model (4.13). And market is cleared by (4.14). In this MOPEC problem, all the decision makers are at the same level, i.e., there is no lower level decision makers but just equality market clearing constraints, so that each of model (4.4), (4.13), (4.14) can be written as VIs and be solved simultaneously.

Proof (Lemma 4.4.3). On one hand, if  $\rho^*$  is a maxinf-point of the Walrasian with  $W(\rho^*, \cdot) \geq 0$ , it

follows that for all unit vectors  $e^s = (0, \dots, 1, \dots)$ , the  $s$ -th entry is 1,  $W(\rho^*, e^s) \geq 0$  which implies  $ES_s(\rho^*) = 0$ . On the other hand, if  $\rho^*$  is an equilibrium prices, i.e.  $ES_s(\rho^*) = 0$ , it follows that  $W(\rho^*, \cdot) = 0$ . Because  $W(\rho, \cdot) \leq 0$  for any  $\rho$ ,  $\rho^*$  is a maxinf point of the Walrasian.

## B.2 Relationship between Waiting Time and Driver/Rider Flow

Equation (4.7) can be better explained graphically in Figure B.1. Line 1 and 2 indicates the accumulative drivers (slope  $f_1$ ) and riders (slope  $f_2$ ) arrival curves, respectively. Line 3 represents the accumulative matching (with slope  $m$  calculated from Cobb-Douglas matching function). Figure B.1 illustrates the case when the arrival rate of drivers  $f_1$  is higher than riders  $f_2$ . We use waiting time for drivers  $t_1$  as an example. The calculation process for waiting time of riders  $t_2$  are similar. The area of the shaded region indicates the total delay of drivers  $T_1$  till time period  $T$ , which can be numerically calculated based on equation (4.7b). Then, the average delay per driver ( $t_1$ ) can be calculated based on equation (4.7c). By varying  $f_1$  and  $f_2$ , we are able to generate a corresponding  $t_1$  numerically, which will be used for estimating parameters in (4.8).

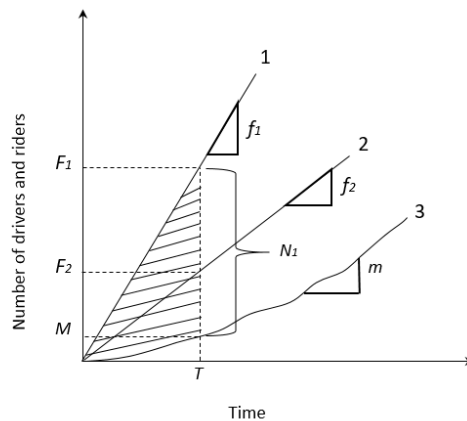


Figure B.1: Graphical Representation of Waiting Time

Following this data generation procedures outlined above, we randomly sample 10,000 pairs of  $f_1$  and  $f_2$  uniformly between  $[0, 100]$  with a maximum difference of 10 and calculate  $t_1$  and  $t_2$ . To use linear regression techniques, we take  $\ln$  of both sides in equation 4.8 and estimate coefficients  $\ln a_0$ ,  $a_1$ , and  $a_2$  in  $\ln t_i \sim \ln a_0 + a_1 \ln f_i + a_2 \ln f_{-i}$ . A sample of data input for regression is summarized in Table B.1. The linear regression output is summarized in Table B.2.

Table B.1: Sample Regression Input

<b>f1</b>	<b>f2</b>	<b>t1</b>	<b>t2</b>	<b>log_f1</b>	<b>log_f2</b>	<b>log_t1</b>	<b>log_t2</b>
10	9	6.03	3.26	2.30	2.20	1.80	1.18
23	15	11.68	1.38	3.14	2.71	2.46	0.32
27	26	4.31	3.29	3.30	3.26	1.46	1.19
39	43	2.45	5.11	3.66	3.76	0.90	1.63
53	50	4.29	2.69	3.97	3.91	1.46	0.99
63	57	4.98	2.24	4.14	4.04	1.61	0.81
66	67	3.05	3.47	4.19	4.20	1.12	1.24
82	76	4.41	2.31	4.41	4.33	1.48	0.84
86	91	2.43	4.00	4.45	4.51	0.89	1.39

### B.3 Extensions

red

#### B.3.1 Extension of Linear Demand Function

We note that the linear demand function (4.2) in Section 4.3.1 is used to keep a clear focus of this paper and it does not restrict the applicability of our modeling and computational approach for more sophisticated demand models. For example, one can adopt nonlinear demand models, such as a logit model, to describe the ride-sourcing demand as a non-linear (decreasing) function of

Table B.2: Regression Summary Output of (4.8)

SUMMARY OUTPUT

<i>Regression Statistics</i>	
Multiple R	0.95784929
R Square	0.91747526
Adjusted R Square	0.91745875
Standard Error	0.04271119
Observations	10000

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	2	202.751098	101.375549	55571.218	0
Residual	9997	18.2369831	0.00182425		
Total	9999	220.988081			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	1.84298058	0.00576574	319.643167	0	1.83167856	1.8542826
log_f1	2.24526358	0.00744279	301.669604	0	2.23067421	2.25985295
log_f2	-2.4035486	0.00748194	-321.24684	0	-2.4182147	-2.3888825

prices. We can choose a utility function of choosing ride-sourcing (with driving as the base option) as given in Equation (B.9).

$$U_s = \beta'_{0,s} - \beta'_2 \rho_s \quad (\text{B.9})$$

where  $\beta'_{0,s}$  and  $\beta'_2$  are the utility coefficients for ride-sourcing attractiveness and ride-sourcing prices at location  $s$ . The logistic demand model given utility function (B.9) can be derived in a closed-form expression as Equation (B.10).

$$d_s = \frac{e^{U_s}}{1 + e^{U_s}} D_s \quad (\text{B.10})$$

We can see that the overall modeling strategy for rider demand is the same for both the linear and nonlinear demand cases, i.e. Equations (4.2) and (B.10). We assume a total of  $D_s$  travelers choosing between ride-sourcing and driving. Prices are one of the key components influencing travelers' mode choices in both cases. Following these two models, within a total of travelers  $D_s$ , we can determine how many travelers choose ride-sourcing (denoted as  $d_s$ ) and driving (denoted as  $D_s - d_s$ ).

When rider utility of choosing ride-sourcing over driving is modeled using a logistic model and matching time can be ignored, i.e., model (B.9), the ride-sourcing demand at each location  $s$ , i.e., Equation (B.10), can be obtained by solving a strictly convex optimization model (B.11),  $\forall s$ .

$$\underset{d_s \in \mathbb{R}_+}{\text{minimize}} \quad [d_s (\ln d_s - 1 + \beta'_2 \rho_s - \beta'_{0,s}) + (D_s - d_s) (\ln(D_s - d_s) - 1)] \quad (\text{B.11})$$

Following similar computational strategies as linear demand case (Section 4.4.1), when equilibrium prices  $\rho$  exist, model (4.5) (replacing linear demand function (4.2) with logistic model (B.10)) can be reformulated as model (B.12), which is a convex optimization problem.

$$\underset{\hat{v}, \hat{v}, q, d \in \mathbb{R}_+}{\text{minimize}} \quad \frac{\beta_1}{\beta_2} \sum_{a \in \mathcal{A}} \int_0^{v_a} t_a(v_a) du + \frac{1}{\beta_2} \sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{S}} q_{rs} (\ln q_{rs} - 1 - \beta_{0,s}) + \frac{1}{\beta_2} \sum_{s \in \mathcal{S}} [d_s (\ln d_s - 1 - \beta'_{0,s}) + \beta'_2 (D_s - d_s) (\ln(D_s - d_s) - 1)] \quad (\text{B.12a})$$

subject to

$$(4.4b \sim 4.4e)$$

$$(\rho_s) \quad \sum_{r \in \mathcal{R}} q_{rs} = d_s, \quad \forall s \quad (\text{B.12b})$$

red

### B.3.2 Extension of Network Effects of Available Drivers

In model (4.4), we treat spatial availability of drivers  $Q_r$  as exogenous variables. Exogenous vehicle fleet size assumption has also been made in Zhou et al. (2021) to study ride-sourcing platform-integration. But the proposed model can also consider the network effects of the rider's destination on the spatial distribution of available drivers, i.e., if a driver takes a rider from  $s$  to  $r$ , then thereafter the driver will be at  $r$ .

$Q_r$  not only depends on how many riders arrive at location  $r$  (denoted as  $d_r$ ), but also depends on the rate of drivers start/end their service during the studied period (denoted as  $Q_r^0 \in$ ). The flow rate of drivers available at location  $r$  can be calculated in Equation (B.13).

$$Q_r = d_r + Q_r^0 \quad (\text{B.13})$$

Denote the proportion of riders going from  $s$  to  $r$  as  $\delta_{sr}$  ( $\sum_{r \in R} \delta_{sr} = 1$ ).  $d_r$  can be calculated in Equation (B.14).

$$d_r = \sum_{s \in S} \delta_{sr} d_s \quad (\text{B.14})$$

There are two approaches to consider the network effects within our modeling framework.

The first approach is an iterative approach. Because the formulation (4.4) and convex reformulation (4.15) are valid for any  $Q_r$ . We can recalculate  $Q_r$  iteratively by replacing  $d_r$  with the values from

the last iteration. The equilibrium  $Q_r$  achieved in this way has incorporated the spatial distribution of riders' drop-off locations. The convergence of  $Q_r$  is shown in Figure B.2 for Sioux Falls network example. We can see that within three iterations, the available drivers flow rate  $Q_r$  is stabilized for every location.

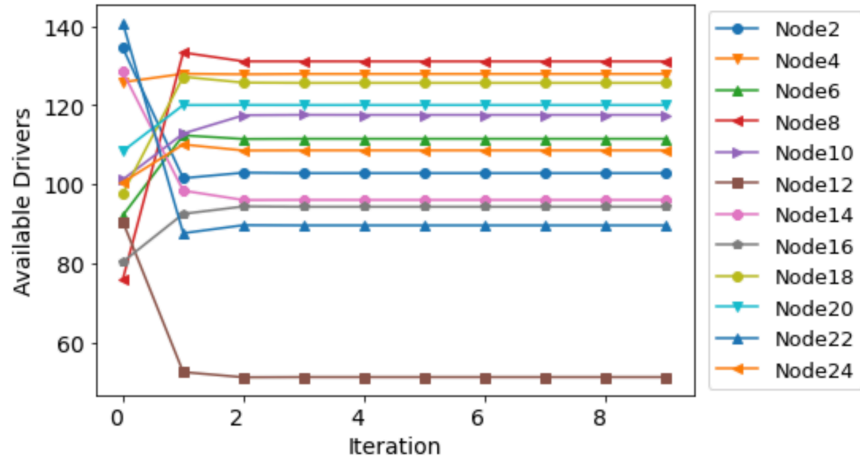


Figure B.2: Convergence of Available Drivers Distribution

The second approach, we can add constraint (B.13) as a global constraint. Since we have proved Lemma 4.4.1 for any given  $Q_r \in \mathbb{R}_+$ , adding constraint (B.13) does not affect the existence and uniqueness of equilibrium prices and the convex reformulation thereafter. In addition, solving model (4.15) with additional global constraint (B.13) has minimum impact on the computational complexity since model (4.15) remains a convex optimization. After we solve for  $Q_r$  that satisfies condition (B.13), we can solve (4.15) separately with the equilibrium  $Q_r$  as input to get the corresponding dual variables  $\rho_s$ . The solution of  $Q_r$  following the second approach, i.e. adding global constraint of (B.13), is identical with the results from the first approach presented in Figure B.2.



## REFERENCES

- Abdel-Aty, M., Dilmore, J., Dhindsa, A., 2006. Evaluation of variable speed limits for real-time freeway safety improvement. *Accident analysis & prevention* 38, 335–345.
- ABI, 2018. *ABI Research Forecasts 8 Million Vehicles to Ship with SAE Level 3, 4 and 5 Autonomous Technology in 2025*. <https://www.abiresearch.com/press/abi-research-forecasts-8-million-vehicles-ship-sae-level-3-4-and-5-autonomous-technology-2025/>[Access ed: 2021-09-30].
- Acemoglu, D., Makhdoumi, A., Malekian, A., Ozdaglar, A., 2018. Informational braess' paradox: The effect of information on traffic congestion. *Operations Research* 66, 893–917.
- Achterberg, T., 2009. Scip: solving constraint integer programs. *Mathematical Programming Computation* 1, 1–41.
- Afifah, F., Guo, Z., Abdel-Aty, M., 2023. System-level impacts of en-route information sharing considering adaptive routing. *Transportation Research Part C: Emerging Technologies* 149, 104075.
- Alemi, F., Circella, G., Handy, S., Mokhtarian, P., 2018. What influences travelers to use uber? exploring the factors affecting the adoption of on-demand ride services in california. *Travel Behaviour and Society* 13, 88–104.
- Allaby, P., Hellinga, B., Bullock, M., 2007. Variable speed limits: Safety and operational impacts of a candidate control strategy for freeway applications. *IEEE Transactions on Intelligent Transportation Systems* 8, 671–680.
- Amini, E., Omidvar, A., Elefteriadou, L., 2021. Optimizing operations at freeway weaves with

- connected and automated vehicles. *Transportation Research Part C: Emerging Technologies* 126, 103072.
- Antoniou, C., Koutsopoulos, H.N., Ben-Akiva, M., Chauhan, A.S., 2011. Evaluation of diversion strategies using dynamic traffic assignment. *Transportation planning and technology* 34, 199–216.
- Baghali, S., Guo, Z., 2021. Impacts of privately owned electric vehicles on distribution system resilience: A multi-agent optimization approach. *arXiv preprint arXiv:2105.03828* .
- Baillon, J.B., Cominetti, R., 2008. Markovian traffic equilibrium. *Mathematical Programming* 111, 33–56.
- Ban, X., Li, Y., et al., 2009. Optimal use of changeable message signs for displaying travel times .
- Ban, X.J., Dessouky, M., Pang, J.S., Fan, R., 2019. A general equilibrium model for transportation systems with e-hailing services and flow congestion. *Transportation Research Part B: Methodological* 129, 273–304.
- Banerjee, S., Johari, R., Riquelme, C., 2015. Pricing in ride-sharing platforms: A queueing-theoretic approach, in: *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, ACM. pp. 639–639.
- Battifarano, M., Qian, Z.S., 2019. Predicting real-time surge pricing of ride-sourcing companies. *Transportation Research Part C: Emerging Technologies* 107, 444–462.
- Beckmann, M., McGuire, C., Winsten, C.B., 1956. *Studies in the Economics of Transportation*. Report.
- Beckmann, M.J., McGuire, C.B., Winsten, C.B., 1955. *Studies in the Economics of Transportation*. Rand Corporation.

- Bel, G., Rosell, J., 2013. Effects of the 80 km/h and variable speed limits on air pollution in the metropolitan area of barcelona. *Transportation Research Part D: Transport and Environment* 23, 90–97.
- Beojone, C.V., Geroliminis, N., 2021. On the inefficiency of ride-sourcing services towards urban congestion. *Transportation research part C: emerging technologies* 124, 102890.
- Bimpikis, K., Candogan, O., Saban, D., 2019. Spatial pricing in ride-sharing networks. *Operations Research* .
- Böhm, M., Frötscher, A., 2009. Bidirectional i2v communication applications for advanced in-vehicle isa systems, in: *2009 IEEE Vehicular Networking Conference (VNC)*, IEEE. pp. 1–6.
- Bonsall, P., 1992. The influence of route guidance advice on route choice in urban networks. *Transportation* 19, 1–23.
- Boukerche, A., Zhong, D., Sun, P., 2021. A novel reinforcement learning-based cooperative traffic signal system through max-pressure control. *IEEE Transactions on Vehicular Technology* 71, 1187–1198.
- Boyles, S.D., Waller, S.T., 2010. A mean-variance model for the minimum cost flow problem with stochastic arc costs. *Networks* 56, 215–227.
- Boyles, S.D., Waller, S.T., 2011. Optimal information location for adaptive routing. *Networks and Spatial Economics* 11, 233–254.
- Branston, D., 1976. Link capacity functions: A review. *Transportation research* 10, 223–236.
- Cachon, G.P., Daniels, K.M., Lobel, R., 2017. The role of surge pricing on a service platform with self-scheduling capacity. *Manufacturing & Service Operations Management* 19, 368–384.

- Calderone, D., Sastry, S.S., 2017. Markov decision process routing games, in: 2017 ACM/IEEE 8th International Conference on Cyber-Physical Systems (ICCPS), IEEE. pp. 273–280.
- Camerer, C., Babcock, L., Loewenstein, G., Thaler, R., 1997. Labor supply of new york city cabdrivers: One day at a time. *The Quarterly Journal of Economics* 112, 407–441.
- Carlson, R.C., Papamichail, I., Papageorgiou, M., 2011. Local feedback-based mainstream traffic flow control on motorways using variable speed limits. *IEEE Transactions on intelligent transportation systems* 12, 1261–1276.
- Carlson, R.C., Papamichail, I., Papageorgiou, M., 2014. Integrated feedback ramp metering and mainstream traffic flow control on motorways using variable speed limits. *Transportation research part C: Emerging technologies* 46, 209–221.
- Casas, N., 2017. Deep deterministic policy gradient for urban traffic light control. *arXiv preprint arXiv:1703.09035* .
- Castiglione, J., Cooper, D., Sana, B., Tischler, D., Chang, T., Erhardt, G.D., Roy, S., Chen, M., Mucci, A., 2018. Tncs & congestion .
- Castillo, J.C., Knoepfle, D., Weyl, G., 2017. Surge pricing solves the wild goose chase, in: *Proceedings of the 2017 ACM Conference on Economics and Computation*, ACM. pp. 241–242.
- Chen, C., Yao, F., Mo, D., Zhu, J., Chen, X.M., 2021. Spatial-temporal pricing for ride-sourcing platform with reinforcement learning. *Transportation Research Part C: Emerging Technologies* 130, 103272.
- Chen, L., Mislove, A., Wilson, C., 2015. Peeking beneath the hood of uber, in: *Proceedings of the 2015 Internet Measurement Conference*, ACM. pp. 495–508.
- Chen, M.K., Sheldon, M., 2016. Dynamic pricing in a labor market: Surge pricing and flexible work on the uber platform., in: *Ec*, p. 455.

- Chen, Y., Li, Y., Xu, D., Xiao, L., 2018. Dqn-based power control for iot transmission against jamming, in: 2018 IEEE 87th Vehicular Technology Conference (VTC Spring), IEEE. pp. 1–5.
- Chen, Z., Liu, X.C., Zhang, G., 2016. Non-recurrent congestion analysis using data-driven spatiotemporal approach for information construction. *Transportation Research Part C: Emerging Technologies* 71, 19–31.
- Chorus, C.G., Molin, E.J., Van Wee, B., 2006. Use and effects of advanced traveller information services (atis): a review of the literature. *Transport Reviews* 26, 127–149.
- Daganzo, C.F., Sheffi, Y., 1977. On stochastic models of traffic assignment. *Transportation science* 11, 253–274.
- Dechenaux, E., Mago, S.D., Razzolini, L., 2014. Traffic congestion: an experimental study of the downs-thomson paradox. *Experimental economics* 17, 461–487.
- Deride, J., Jofré, A., Wets, R.J., 2019. Solving deterministic and stochastic equilibrium problems via augmented walrasian. *Computational Economics* 53, 315–342.
- Desjardins, C., Chaib-Draa, B., 2011. Cooperative adaptive cruise control: A reinforcement learning approach. *IEEE Transactions on intelligent transportation systems* 12, 1248–1260.
- Du, L., Han, L., Chen, S., 2015. Coordinated online in-vehicle routing balancing user optimality and system optimality through information perturbation. *Transportation Research Part B: Methodological* 79, 121–133.
- Du, L., Han, L., Li, X.Y., 2014. Distributed coordinated in-vehicle online routing using mixed-strategy congestion game. *Transportation Research Part B: Methodological* 67, 1–17.
- Du, Z., HomChaudhuri, B., Pisu, P., 2018. Hierarchical distributed coordination strategy of connected and automated vehicles at multiple intersections. *Journal of Intelligent Transportation Systems* 22, 144–158.

- Dzisi, E.K., Ackaah, W., Aprimah, B.A., Adjei, E., 2020. Understanding demographics of ride-sourcing and the factors that underlie its use among young people. *Scientific African* 7, e00288.
- El-Tantawy, S., Abdulhai, B., Abdelgawad, H., 2013. Multiagent reinforcement learning for integrated network of adaptive traffic signal controllers (marlin-atssc): methodology and large-scale application on downtown toronto. *IEEE Transactions on Intelligent Transportation Systems* 14, 1140–1150.
- Elliott, D., Keen, W., Miao, L., 2019. Recent advances in connected and automated vehicles. *Journal of Traffic and Transportation Engineering (English Edition)* 6, 109–131.
- Emmerink, R.H., Axhausen, K.W., Nijkamp, P., Rietveld, P., 1995. Effects of information in road transport networks with recurrent congestion. *Transportation* 22, 21–53.
- Erdmann, J., 2014. Lane-changing model in sumo. *Proceedings of the SUMO2014 modeling mobility with open data* 24, 77–88.
- Evans, S.P., 1976. Derivation and analysis of some models for combining trip distribution and assignment. *Transportation research* 10, 37–57.
- Fagnant, D.J., Kockelman, K., 2015. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice* 77, 167–181.
- Farber, H.S., 2005. Is tomorrow another day? the labor supply of new york city cabdrivers. *Journal of political Economy* 113, 46–82.
- Farrag, S., El-Hansali, M.Y., Yasar, A., Shakshuki, E.M., 2020. Simulation-based evaluation of using variable speed limit in traffic incidents. *Procedia Computer Science* 175, 340–348.
- Ferris, M.C., Wets, R., 2013. Mopec: multiple optimization problems with equilibrium constraints.

- Fisk, C., Nguyen, S., 1982. Solution algorithms for network equilibrium models with asymmetric user costs. *Transportation Science* 16, 361–381.
- Florian, M., Nguyen, S., Ferland, J., 1975. On the combined distribution-assignment of traffic. *Transportation Science* 9, 43–53.
- Frejo, J.R.D., Camacho, E.F., 2012. Global versus local mpc algorithms in freeway traffic control with ramp metering and variable speed limits. *IEEE Transactions on intelligent transportation systems* 13, 1556–1565.
- Friesz, T.L., Han, K., 2019. The mathematical foundations of dynamic user equilibrium. *Transportation research part B: methodological* 126, 309–328.
- Fujimoto, S., Hoof, H., Meger, D., 2018. Addressing function approximation error in actor-critic methods, in: *International conference on machine learning*, PMLR. pp. 1587–1596.
- Fyfe, M., Sayed, T., 2017. Safety evaluation of connected vehicles for a cumulative travel time adaptive signal control microsimulation using the surrogate safety assessment model. *Technical Report*.
- Gao, S., 2012. Modeling strategic route choice and real-time information impacts in stochastic and time-dependent networks. *IEEE Transactions on Intelligent Transportation Systems* 13, 1298–1311.
- Gao, S., Chabini, I., 2006. Optimal routing policy problems in stochastic time-dependent networks. *Transportation Research Part B: Methodological* 40, 93–122.
- Gao, S., Huang, H., 2012. Real-time traveler information for optimal adaptive routing in stochastic time-dependent networks. *Transportation Research Part C: Emerging Technologies* 21, 196–213.

- Gregurić, M., Kušić, K., Ivanjko, E., 2022. Impact of deep reinforcement learning on variable speed limit strategies in connected vehicles environments. *Engineering Applications of Artificial Intelligence* 112, 104850.
- Gregurić, M., Vujić, M., Alexopoulos, C., Miletić, M., 2020. Application of deep reinforcement learning in traffic signal control: An overview and impact of open traffic data. *Applied Sciences* 10, 4011.
- Grumert, E., Ma, X., Tapani, A., 2015. Analysis of a cooperative variable speed limit system using microscopic traffic simulation. *Transportation research part C: emerging technologies* 52, 173–186.
- Guda, H., Subramanian, U., 2019. Your uber is arriving: Managing on-demand workers through surge pricing, forecast communication, and worker incentives. *Management Science* 65, 1995–2014.
- Guériau, M., Dusparic, I., 2018. Samod: Shared autonomous mobility-on-demand using decentralized reinforcement learning, in: *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, IEEE. pp. 1558–1563.
- Guo, Z., Afifah, F., Qi, J., Baghali, S., 2021. A stochastic multiagent optimization framework for interdependent transportation and power system analyses. *IEEE Transactions on Transportation Electrification* 7, 1088–1098.
- Guo, Z., Chen, R.L.Y., Fan, N., Watson, J.P., 2016a. Contingency-constrained unit commitment with intervening time for system adjustments. *IEEE Transactions on Power Systems* 32, 3049–3059.
- Guo, Z., Deride, J., Fan, Y., 2016b. Infrastructure planning for fast charging stations in a competitive market. *Transportation Research Part C: Emerging Technologies* 68, 215–227.



- Guo, Z., Fan, Y., 2017. A stochastic multi-agent optimization model for energy infrastructure planning under uncertainty in an oligopolistic market. *Networks and Spatial Economics* 17, 581–609.
- Gurvich, I., Lariviere, M., Moreno, A., 2019. Operations in the on-demand economy: Staffing services with self-scheduling capacity, in: *Sharing Economy*. Springer, pp. 249–278.
- Hall, R.W., 1996. Route choice and advanced traveler information systems on a capacitated and dynamic network. *Transportation Research Part C: Emerging Technologies* 4, 289–306.
- Han, Y., Hegyi, A., Yuan, Y., Hoogendoorn, S., Papageorgiou, M., Roncoli, C., 2017. Resolving freeway jam waves by discrete first-order model-based predictive control of variable speed limits. *Transportation Research Part C: Emerging Technologies* 77, 405–420.
- Hart, W.E., Laird, C.D., Watson, J.P., Woodruff, D.L., Hackebeil, G.A., Nicholson, B.L., Siirola, J.D., 2017. *Pyomo—optimization modeling in python*. volume 67. Second ed., Springer Science & Business Media.
- Hasibur Rahman, M., Abdel-Aty, M., 2021. Application of connected and automated vehicles in a large-scale network by considering vehicle-to-vehicle and vehicle-to-infrastructure technology. *Transportation Research Record: Journal of the Transportation Research Board* 2675.
- He, F., Wang, X., Lin, X., Tang, X., 2018. Pricing and penalty/compensation strategies of a taxi-hailing platform. *Transportation Research Part C: Emerging Technologies* 86, 263–279.
- Hegyi, A., De Schutter, B., Hellendoorn, H., 2005a. Model predictive control for optimal coordination of ramp metering and variable speed limits. *Transportation Research Part C: Emerging Technologies* 13, 185–209.
- Hegyi, A., De Schutter, B., Hellendoorn, H., Van Den Boom, T., 2002. Optimal coordination

- of ramp metering and variable speed control-an mpc approach, in: Proceedings of the 2002 American Control Conference (IEEE Cat. No. CH37301), IEEE. pp. 3600–3605.
- Hegyi, A., De Schutter, B., Hellendoorn, J., 2005b. Optimal coordination of variable speed limits to suppress shock waves. *IEEE Transactions on intelligent transportation systems* 6, 102–112.
- Hegyi, A., Hoogendoorn, S.P., Schreuder, M., Stoelhorst, H., Viti, F., 2008. Specialist: A dynamic speed limit control algorithm based on shock wave theory, in: 2008 11th international ieee conference on intelligent transportation systems, IEEE. pp. 827–832.
- Hellinga, B., Mandelzys, M., 2011. Impact of driver compliance on the safety and operational impacts of freeway variable speed limit systems. *Journal of transportation engineering* 137, 260–268.
- Henn, V., Ottomanelli, M., 2006. Handling uncertainty in route choice models: From probabilistic to possibilistic approaches. *European journal of operational research* 175, 1526–1538.
- Hoogendoorn, S., Daamen, W., Hoogendoorn, R., Goemans, J., 2013. Assessment of dynamic speed limits on freeway a20 near rotterdam, netherlands. *Transportation research record* 2380, 61–71.
- Hu, B., Hu, M., Zhu, H., 2021. Surge pricing and two-sided temporal responses in ride hailing. *Manufacturing & Service Operations Management* .
- Hydén, C., 1996. Traffic conflicts technique: state-of-the-art. *Traffic safety work with video processing* 37, 3–14.
- Jordanidou, G.R., Roncoli, C., Papamichail, I., Papageorgiou, M., 2014. Feedback-based mainstream traffic flow control for multiple bottlenecks on motorways. *IEEE Transactions on Intelligent Transportation Systems* 16, 610–621.

- Islam, M.T., Hadiuzzaman, M., Fang, J., Qiu, T.Z., El-Basyouny, K., 2013. Assessing mobility and safety impacts of a variable speed limit control strategy. *Transportation research record* 2364, 1–11.
- Jesus, J.C., Bottega, J.A., Cuadros, M.A., Gamarra, D.F., 2019. Deep deterministic policy gradient for navigation of mobile robots in simulated environments, in: 2019 19th International Conference on Advanced Robotics (ICAR), IEEE. pp. 362–367.
- Jin, H.Y., Jin, W.L., 2015. Control of a lane-drop bottleneck through variable speed limits. *Transportation Research Part C: Emerging Technologies* 58, 568–584.
- Jinnai, Y., Park, J.W., Machado, M.C., Konidaris, G., 2019. Exploration in reinforcement learning with deep covering options, in: International Conference on Learning Representations.
- Jofré, A.e, A., Wets, R.J.B., 2009. Variational convergence of bivariate functions: Lopsided convergence. *Mathematical Programming* 116, 275–295.
- Karamanis, R., Angeloudis, P., Sivakumar, A., Stettler, M., 2018. Dynamic pricing in one-sided autonomous ride-sourcing markets, in: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), IEEE. pp. 3645–3650.
- Konda, V., Tsitsiklis, J., 1999. Actor-critic algorithms. *Advances in neural information processing systems* 12.
- Kušić, K., Dusparic, I., Guériau, M., Gregurić, M., Ivanjko, E., 2020a. Extended variable speed limit control using multi-agent reinforcement learning, in: 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), IEEE. pp. 1–8.
- Kušić, K., Ivanjko, E., Gregurić, M., Miletić, M., 2020b. An overview of reinforcement learning methods for variable speed limit control. *Applied Sciences* 10, 4917.

- Kušić, K., Ivanjko, E., Vrbanić, F., Gregurić, M., Dusparic, I., 2021. Spatial-temporal traffic flow control on motorways using distributed multi-agent reinforcement learning. *Mathematics* 9, 3081.
- Kwon, J., Mauch, M., Varaiya, P., 2006. Components of congestion: Delay from incidents, special events, lane closures, weather, potential ramp metering gain, and excess demand. *Transportation Research Record* 1959, 84–91.
- Lagos, R., 2000. An alternative approach to search frictions. *Journal of Political Economy* 108, 851–873.
- Lam, W.H., Chan, K., 1996. A stochastic traffic assignment model for road network with travel time information via variable message signs, in: *Proceedings of Conference on Intelligent Vehicles*, IEEE. pp. 99–104.
- Lee, C., Hellinga, B., Saccomanno, F., 2003. Real-time crash prediction model for application to crash prevention in freeway traffic. *Transportation Research Record* 1840, 67–77.
- Lee, C., Hellinga, B., Saccomanno, F., 2004. Assessing safety benefits of variable speed limits. *Transportation Research Record* 1897, 183–190.
- Lee, J., Park, B., 2012. Development and evaluation of a cooperative vehicle intersection control algorithm under the connected vehicles environment. *IEEE transactions on intelligent transportation systems* 13, 81–90.
- Levinson, D., 2003. The value of advanced traveler information systems for route choice. *Transportation Research Part C: Emerging Technologies* 11, 75–87.
- Li, D., Ranjitkar, P., 2015. A fuzzy logic-based variable speed limit controller. *Journal of advanced transportation* 49, 913–927.

- Li, D., Ranjitkar, P., Ceder, A., 2014. A logic tree based algorithm for variable speed limit controllers to manage recurrently congested bottlenecks. Technical Report.
- Li, J., Yu, T., Zhang, X., Li, F., Lin, D., Zhu, H., 2021a. Efficient experience replay based deep deterministic policy gradient for agc dispatch in integrated energy system. *Applied Energy* 285, 116386.
- Li, M., Lin, X., He, F., Jiang, H., 2016. Optimal locations and travel time display for variable message signs. *Transportation Research Part C: Emerging Technologies* 69, 418–435.
- Li, S., Yang, H., Poolla, K., Varaiya, P., 2021b. Spatial pricing in ride-sourcing markets under a congestion charge. *Transportation Research Part B: Methodological* 152, 18–45.
- Li, X., Ke, J., Yang, H., Wang, H., Zhou, Y., 2021c. A general matching function for ride-sourcing services. Available at SSRN 3915450 .
- Li, Z., Liu, P., Xu, C., Duan, H., Wang, W., 2017. Reinforcement learning-based variable speed limit control strategy to reduce traffic congestion at freeway recurrent bottlenecks. *IEEE transactions on intelligent transportation systems* 18, 3204–3217.
- Liang, Y., Guo, C., Ding, Z., Hua, H., 2020. Agent-based modeling in electricity market using deep deterministic policy gradient algorithm. *IEEE transactions on power systems* 35, 4180–4192.
- Lillicrap, T.P., Hunt, J.J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., Wierstra, D., 2015. Continuous control with deep reinforcement learning. arXiv preprint arXiv:1509.02971 .
- Lin, D.Y., Eluru, N., Waller, S.T., Bhat, C.R., 2008. Integration of activity-based modeling and dynamic traffic assignment. *Transportation Research Record* 2076, 52–61.
- Lin, P.W., Kang, K.P., Chang, G.L., 2004. Exploring the effectiveness of variable speed limit controls on highway work-zone operations, in: *Intelligent transportation systems*, Taylor & Francis. pp. 155–168.

- Lindsey, R., Daniel, T., Gisches, E., Rapoport, A., 2014. Pre-trip information and route-choice decisions with stochastic travel conditions: Theory. *Transportation Research Part B: Methodological* 67, 187–207.
- Liu, P., Fan, W.D., 2021. Exploring the impact of connected and autonomous vehicles on mobility and environment at signalized intersections through vehicle-to-infrastructure (v2i) and infrastructure-to-vehicle (i2v) communications. *Transportation planning and technology* 44, 129–138.
- Liu, P., Liu, Y., 2018. Optimal information provision at bottleneck equilibrium with risk-averse travelers. *Transportation Research Record* 2672, 69–78.
- Lu, X., Gao, S., Ben-Elia, E., 2011. Information impacts on route choice and learning behavior in a congested network: experimental approach. *Transportation Research Record* 2243, 89–98.
- Ma, F., Yang, Y., Wang, J., Li, X., Wu, G., Zhao, Y., Wu, L., Aksun-Guvenc, B., Guvenc, L., 2021. Eco-driving-based cooperative adaptive cruise control of connected vehicles platoon at signalized intersections. *Transportation Research Part D: Transport and Environment* 92, 102746.
- Ma, J., Smith, B.L., Zhou, X., 2016. Personalized real-time traffic information provision: Agent-based optimization model and solution framework. *Transportation Research Part C: Emerging Technologies* 64, 164–182.
- Mahmassani, H.S., 2001. Dynamic network traffic assignment and simulation methodology for advanced system management applications. *Networks and spatial economics* 1, 267–292.
- McNally, M.G., 2007. *The four-step model*. Emerald Group Publishing Limited.
- Morando, M.M., Tian, Q., Truong, L.T., Vu, H.L., 2018. Studying the safety impact of autonomous vehicles using simulation-based surrogate safety measures. *Journal of Advanced Transportation* 2018.

- Morgan, J., Orzen, H., Sefton, M., 2009. Network architecture and traffic flows: Experiments on the pigou–knight–downs and braess paradoxes. *Games and economic behavior* 66, 348–372.
- Müller, E.R., Carlson, R.C., Kraus, W., Papageorgiou, M., 2015. Microsimulation analysis of practical aspects of traffic control with variable speed limits. *IEEE Transactions on Intelligent Transportation Systems* 16, 512–523.
- Nagurney, A., Siokos, S., 1997. Variational inequalities, in: *Financial Networks*. Springer, pp. 49–73.
- NHTSA, 2023. *The Economic and Societal Impact of Motor Vehicle Crashes, 2019 (Revised)*. Technical Report DOT HS 813 003. National Center for Statistics and Analysis, U.S. Department of Transportation. URL: <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813003>.
- Nie, Y.M., 2017. How can the taxi industry survive the tide of ridesourcing? evidence from shenzhen, china. *Transportation Research Part C: Emerging Technologies* 79, 242–256.
- Nourinejad, M., Ramezani, M., 2020. Ride-sourcing modeling and pricing in non-equilibrium two-sided markets. *Transportation Research Part B: Methodological* 132, 340–357.
- NYCTLC, 2019. *New York State’s Congestion Surcharge*. <https://www1.nyc.gov/site/tlc/about/congestion-surcharge.page>[Accessed: 2022-02-15].
- Papadoulis, A., Quddus, M., Imprialou, M., 2018. Estimating the Corridor-Level Safety Impact of Connected and Autonomous Vehicles. Technical Report.
- Papadoulis, A., Quddus, M., Imprialou, M., 2019. Evaluating the safety impact of connected and autonomous vehicles on motorways. *Accident Analysis & Prevention* 124, 12–22.
- Papageorgiou, M., Kosmatopoulos, E., Papamichail, I., 2008. Effects of variable speed limits on motorway traffic flow. *Transportation Research Record* 2047, 37–48.

- Park, J.J., Lee, Y.M., Park, J.B., Kang, J.G., 2008. The effect of point to point speed enforcement systems on traffic flow characteristics. *Journal of Korean Society of Transportation* 26, 85–95.
- Peng, C., Xu, C., 2021. Combined variable speed limit and lane change guidance for secondary crash prevention using distributed deep reinforcement learning. *Journal of Transportation Safety & Security* , 1–26.
- Pérez, J., Seco, F., Milanés, V., Jiménez, A., Díaz, J.C., De Pedro, T., 2010. An rfid-based intelligent vehicle speed controller using active traffic signals. *Sensors* 10, 5872–5887.
- Philpott, A., Ferris, M., Wets, R., 2016. Equilibrium, uncertainty and risk in hydro-thermal electricity systems. *Mathematical Programming* 157, 483–513.
- Pierog, K., 2019. Chicago approves traffic congestion tax on ride-hailing services. reuters, 2019.
- Powell, M., 2009. The bobyqa algorithm for bound constrained optimization without derivatives. Cambridge NA Report NA2009/06, University of Cambridge, Cambridge .
- Qian, X., Lei, T., Xue, J., Lei, Z., Ukkusuri, S.V., 2020. Impact of transportation network companies on urban congestion: Evidence from large-scale trajectory data. *Sustainable Cities and Society* 55, 102053.
- Qiu, C., Hu, Y., Chen, Y., Zeng, B., 2019. Deep deterministic policy gradient (ddpg)-based energy harvesting wireless communications. *IEEE Internet of Things Journal* 6, 8577–8588.
- Qu, X., Yu, Y., Zhou, M., Lin, C.T., Wang, X., 2020. Jointly dampening traffic oscillations and improving energy consumption with electric, connected and automated vehicles: a reinforcement learning based approach. *Applied Energy* 257, 114030.
- Rahman, M.S., Abdel-Aty, M., Lee, J., Rahman, M.H., 2019. Safety benefits of arterials' crash risk under connected and automated vehicles. *Transportation Research Part C: Emerging Technologies* 100, 354–371.



- Rambha, T., Boyles, S.D., Unnikrishnan, A., Stone, P., 2018. Marginal cost pricing for system optimal traffic assignment with recourse under supply-side uncertainty. *Transportation Research Part B: Methodological* 110, 104–121.
- Rapoport, A., Kugler, T., Dugar, S., Gisches, E.J., 2009. Choice of routes in congested traffic networks: Experimental tests of the braess paradox. *Games and Economic Behavior* 65, 538–571.
- Rochet, J.C., Tirole, J., 2006. Two-sided markets: a progress report. *The RAND journal of economics* 37, 645–667.
- Rockafellar, R.T., Wets, R.B., 1976. Nonanticipativity and 1 1-martingales in stochastic optimization problems, in: *Stochastic Systems: Modeling, Identification and Optimization, II*. Springer, pp. 170–187.
- Rysman, M., 2009. The economics of two-sided markets. *Journal of economic perspectives* 23, 125–43.
- SAE international, 2018. Taxonomy and definitions for terms related to shared mobility and enabling technologies. SAE International,(J3163) .
- Schmidt-Dumont, T., Van Vuuren, J., 2019. A case for the adoption of decentralised reinforcement learning for the control of traffic flow on south african highways. *Journal of the South African Institution of Civil Engineering* 61, 7–19.
- Sheffi, Y., 1985. Urban transportation networks: equilibrium analysis with mathematical programming methods .
- Shi, Q., Abdel-Aty, M., 2015. Big data applications in real-time traffic operation and safety monitoring and improvement on urban expressways. *Transportation Research Part C: Emerging Technologies* 58, 380–394.

- Silver, D., Lever, G., Heess, N., Degris, T., Wierstra, D., Riedmiller, M., 2014. Deterministic policy gradient algorithms, in: International conference on machine learning, PMLR. pp. 387–395.
- Sinha, A., Malo, P., Deb, K., 2017. A review on bilevel optimization: From classical to evolutionary approaches and applications. *IEEE Transactions on Evolutionary Computation* 22, 276–295.
- Sukuvaara, T., Nurmi, P., 2009. Wireless traffic service platform for combined vehicle-to-vehicle and vehicle-to-infrastructure communications. *IEEE Wireless Communications* 16, 54–61.
- Sundaram, S., Koutsopoulos, H.N., Ben-Akiva, M., Antoniou, C., Balakrishna, R., 2011. Simulation-based dynamic traffic assignment for short-term planning applications. *Simulation Modelling Practice and Theory* 19, 450–462.
- Sutton, R.S., Barto, A.G., 1999. Reinforcement learning: An introduction. *Robotica* 17, 229–235.
- Sutton, R.S., McAllester, D., Singh, S., Mansour, Y., 1999. Policy gradient methods for reinforcement learning with function approximation. *Advances in neural information processing systems* 12.
- Tang, W., Wang, H., Wang, Y., Chen, C., Chen, X.M., 2022. A bi-level optimization model for ride-sourcing platform’s spatial pricing strategy. *Journal of Advanced Transportation* 2022.
- Taylor, T.A., 2018. On-demand service platforms. *Manufacturing & Service Operations Management* 20, 704–720.
- Tesauro, G., et al., 1995. Temporal difference learning and td-gammon. *Communications of the ACM* 38, 58–68.
- Toledo, T., Beinhaker, R., 2006. Evaluation of the potential benefits of advanced traveler information systems. *Journal of intelligent transportation systems* 10, 173–183.

- Treiber, M., Kesting, A., 2017. The intelligent driver model with stochasticity-new insights into traffic flow oscillations. *Transportation research procedia* 23, 174–187.
- Tsirimpa, A., Polydoropoulou, A., Antoniou, C., 2007. Development of a mixed multi-nomial logit model to capture the impact of information systems on travelers' switching behavior. *Journal of Intelligent Transportation Systems* 11, 79–89.
- TTI, 2021. 2021 urban mobility report. URL: <https://static.tti.tamu.edu/tti.tamu.edu/documents/mobility-report-2021.pdf>.
- Unnikrishnan, A., Waller, S.T., 2009. User equilibrium with recourse. *Networks and Spatial Economics* 9, 575.
- Vignon, D.A., Yin, Y., Ke, J., 2021. Regulating ridesourcing services with product differentiation and congestion externality. *Transportation Research Part C: Emerging Technologies* 127, 103088.
- Vrbanić, F., Ivanjko, E., Mandžuka, S., Miletić, M., 2021. Reinforcement learning based variable speed limit control for mixed traffic flows, in: *2021 29th Mediterranean Conference on Control and Automation (MED)*, IEEE. pp. 560–565.
- Wächter, A., 2009. Short tutorial: getting started with ipopt in 90 minutes, in: *Dagstuhl Seminar Proceedings*, Schloss Dagstuhl-Leibniz-Zentrum für Informatik.
- Walraven, E., Spaan, M.T., Bakker, B., 2016. Traffic flow optimization: A reinforcement learning approach. *Engineering Applications of Artificial Intelligence* 52, 203–212.
- Wang, C., Xu, Y., Zhang, J., Ran, B., 2022. Integrated traffic control for freeway recurrent bottleneck based on deep reinforcement learning. *IEEE Transactions on Intelligent Transportation Systems* .

- Wang, C., Zhang, J., Xu, L., Li, L., Ran, B., 2019a. A new solution for freeway congestion: Cooperative speed limit control using distributed reinforcement learning. *IEEE Access* 7, 41947–41957.
- Wang, P., Li, H., Chan, C.Y., 2019b. Continuous control for automated lane change behavior based on deep deterministic policy gradient algorithm, in: *2019 IEEE Intelligent Vehicles Symposium (IV)*, IEEE. pp. 1454–1460.
- Wang, Y., Zhang, Y., Hu, J., Li, L., 2012. Using variable speed limits to eliminate wide moving jams: a study based on three-phase traffic theory. *International Journal of Modern Physics C* 23, 1250060.
- Wang, Z., Wu, G., Barth, M.J., 2019c. Cooperative eco-driving at signalized intersections in a partially connected and automated vehicle environment. *IEEE Transactions on Intelligent Transportation Systems* 21, 2029–2038.
- Wardrop, J., 1952a. Some theoretical aspects of road traffic research. *Proceedings of the Institute of Civil Engineers* , 325–378.
- Wardrop, J.G., 1952b. Some theoretical aspects of road traffic research, in: *ICE Proceedings: Engineering Divisions*, Thomas Telford. pp. 325–362.
- Wijayarathna, K.P., Dixit, V.V., Denant-Boemont, L., Waller, S.T., 2017. An experimental study of the online information paradox: Does en-route information improve road network performance? *PLoS One* 12, e0184191.
- Wilson, A.G., 1969. The use of entropy maximising models, in the theory of trip distribution, mode split and route split. *Journal of Transport Economics and Policy* , 108–126.
- Wu, J., Wei, Z., Liu, K., Quan, Z., Li, Y., 2020a. Battery-involved energy management for hy-

- brid electric bus based on expert-assistance deep deterministic policy gradient algorithm. *IEEE Transactions on Vehicular Technology* 69, 12786–12796.
- Wu, T., Zhou, P., Liu, K., Yuan, Y., Wang, X., Huang, H., Wu, D.O., 2020b. Multi-agent deep reinforcement learning for urban traffic light control in vehicular networks. *IEEE Transactions on Vehicular Technology* 69, 8243–8256.
- Wu, Y., Tan, H., Qin, L., Ran, B., 2020c. Differential variable speed limits control for freeway recurrent bottlenecks via deep actor-critic algorithm. *Transportation research part C: emerging technologies* 117, 102649.
- Xie, Y., Zhang, H., Gartner, N.H., Arsava, T., 2017. Collaborative merging strategy for freeway ramp operations in a connected and autonomous vehicles environment. *Journal of Intelligent Transportation Systems* 21, 136–147.
- Xu, J., Hou, Z., Wang, W., Xu, B., Zhang, K., Chen, K., 2018. Feedback deep deterministic policy gradient with fuzzy reward for robotic multiple peg-in-hole assembly tasks. *IEEE Transactions on Industrial Informatics* 15, 1658–1667.
- Yang, H., 1998. Multiple equilibrium behaviors and advanced traveler information systems with endogenous market penetration. *Transportation Research Part B: Methodological* 32, 205–218.
- Yang, H., H. Bell, M.G., 1998. Models and algorithms for road network design: a review and some new developments. *Transport Reviews* 18, 257–278.
- Yang, H., Kitamura, R., Jovanis, P.P., Vaughn, K.M., Abdel-Aty, M.A., 1993. Exploration of route choice behavior with advanced traveler information using neural network concepts. *Transportation* 20, 199–223.
- Yang, H., Leung, C.W., Wong, S., Bell, M.G., 2010. Equilibria of bilateral taxi–customer searching and meeting on networks. *Transportation Research Part B: Methodological* 44, 1067–1083.

- Yang, H., Qin, X., Ke, J., Ye, J., 2020. Optimizing matching time interval and matching radius in on-demand ride-sourcing markets. *Transportation Research Part B: Methodological* 131, 84–105.
- Yang, H., Wong, S., 1998. A network model of urban taxi services. *Transportation Research Part B: Methodological* 32, 235–246.
- Yang, H., Yang, T., 2011. Equilibrium properties of taxi markets with search frictions. *Transportation Research Part B: Methodological* 45, 696–713.
- Yue, L., Abdel-Aty, M., Wu, Y., Wang, L., 2018. Assessment of the safety benefits of vehicles' advanced driver assistance, connectivity and low level automation systems. *Accident Analysis & Prevention* 117, 55–64.
- Zeynivand, A., Javadpour, A., Bolouki, S., Sangaiah, A., Ja'fari, F., Pinto, P., Zhang, W., 2022. Traffic flow control using multi-agent reinforcement learning. *Journal of Network and Computer Applications* 207, 103497.
- Zha, L., Yin, Y., Du, Y., 2018a. Surge pricing and labor supply in the ride-sourcing market. *Transportation Research Part B: Methodological* 117, 708 – 722. URL: <http://www.sciencedirect.com/science/article/pii/S0191261517307683>, doi:<https://doi.org/10.1016/j.trb.2017.09.010>. tRB:ISTTT-22.
- Zha, L., Yin, Y., Xu, Z., 2018b. Geometric matching and spatial pricing in ride-sourcing markets. *Transportation Research Part C: Emerging Technologies* 92, 58–75.
- Zha, L., Yin, Y., Yang, H., 2016. Economic analysis of ride-sourcing markets. *Transportation Research Part C: Emerging Technologies* 71, 249–266.
- Zhang, W., Gai, J., Zhang, Z., Tang, L., Liao, Q., Ding, Y., 2019. Double-dqn based path smooth-

- ing and tracking control method for robotic vehicle navigation. *Computers and Electronics in Agriculture* 166, 104985.
- Zhao, Z., Wu, G., Barth, M., 2021. Corridor-wise eco-friendly cooperative ramp management system for connected and automated vehicles. *Sustainability* 13, 8557.
- Zheng, Y., Ran, B., Qu, X., Zhang, J., Lin, Y., 2019. Cooperative lane changing strategies to improve traffic operation and safety nearby freeway off-ramps in a connected and automated vehicles environment. *IEEE Transactions on Intelligent Transportation Systems* .
- Zhou, M., Qu, X., 2016. Microscopic car-following model for autonomous vehicles using reinforcement learning, in: *Symposium on Innovations in Traffic Flow Theory and Characteristics and TFT Midyear Meeting*, p. 3.
- Zhou, Y., Yang, H., Ke, J., Wang, H., Li, X., 2021. Competition and third-party platform-integration in ride-sourcing markets. *Transportation Research Part B: Methodological* .
- Zhu, F., Ukkusuri, S.V., 2014. Accounting for dynamic speed limit control in a stochastic traffic environment: A reinforcement learning approach. *Transportation research part C: emerging technologies* 41, 30–47.
- Zhu, Z., Hu, C., Zhu, C., Zhu, Y., Sheng, Y., 2021. An improved dueling deep double-q network based on prioritized experience replay for path planning of unmanned surface vehicles. *Journal of Marine Science and Engineering* 9, 1267.
- Zuniga-Garcia, N., Tec, M., Scott, J.G., Ruiz-Juri, N., Machemehl, R.B., 2020. Evaluation of ride-sourcing search frictions and driver productivity: A spatial denoising approach. *Transportation Research Part C: Emerging Technologies* 110, 346–367.