Design and Analysis of Mobility Permit-based Traffic Management Schemes

by

Javad Lessan

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Civil Engineering

Waterloo, Ontario, Canada, 2020

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners. I understand that my thesis may be made electronically available to the public.

Statement of Contributions

I am the sole author of Chapters 1, 3 and 6. Chapters 2 and 5 of this thesis mainly consist of two papers that were co-authored by myself and my supervisor Dr. Liping Fu. Chapters 3 and 4 of this thesis mainly consist of two papers that was co-authored by myself, my supervisor Dr. Liping Fu, and Dr. Chris Bachmann. This thesis includes materials from articles as listed below.

Chapter 2:

• Lessan, J., & Fu, L. (2019). Credit-and permit-based travel demand management state-of-the-art methodological advances. Transportmetrica A: Transport Science, 1-24.

Chapter 3 & 4:

• Lessan, J., Fu, L., & Bachmann, C. (2019, July). A Mechanism for Pricing and Allocation of Mobility Permits on Single-Bottleneck Roadways. In 2019 5th International Conference on Transportation Information and Safety (ICTIS) (pp. 1074-1080). IEEE.

• Lessan, J., Fu, L., & Bachmann, C. (Submitted 2020). Towards Managing Mobility on Road Networks using Permit-based Schemes. Submitted to: Transportation Research Part E: Logistics and Transportation Review.

Chapter 5:

• Lessan, J., & Fu, L. (Submitted 2019). A Mobility Credit Scheme for Managing Traffic on Transportation Networks with Advanced Connectivity Technologies. Submitted to: Journal of Intelligent Transportation Systems: Technology, Planning, and Operations, Revised February 2020.

Abstract

High demand for mobility has undeniably been causing numerous negative impacts on the economy, the society and the environment. As a potential solution to address this challenge, a rapid transition is taking place in the transportation sector with emerging concepts of mobility marketplace. The basic premise is to treat the transportation system and its use as a collection of commodities or services that can be bought from the transportation market. This concept is increasingly becoming a reality with the technological developments in automotive industry such as connected and autonomous vehicles (CAVs). However, there are many policy, design and operation related issues that must be addressed before these traffic management schemes become reality. This thesis research aims at addressing some of these challenges and issues with a specific focus on the two most promising market-driven instruments, namely, mobility permits (MP)- and mobility credits (MC)-based traffic management schemes, which have been proposed to manage travel demand and mitigate traffic congestion by controlling roadway-use right. This research has made several distinctive contributions into the literature. We first conduct a critical review of the state-of-the-art methodological advances on MP- and MC-based travel demand management schemes. We synthesize the relevant body of literature with an in-depth discussion on related studies to provide an improved understanding of the fundamental constructs of these problems, including problem variants, methodologies, and modeling attributes. We also discuss the research gaps and challenges and suggest some possible perspectives and directions for future research.

Based on the gaps identified in the literature review, an integrated framework is proposed for implementing various roadway-use right-based traffic management programs such as MP and MC-based schemes. This framework entails a unique construct for integrating the needs of multiple stakeholders (e.g., road users and authorities), diverse network conditions, and traffic control methods. It allows easy incorporation of different components required for implementing a coordinative mobility scheme, taking into account the influence of the participating players and the underlying issues. The framework can be served as a road-map to future studies on different roadway-use right-based solutions for traffic congestion management.

With our proposed framework, we then focus on addressing various specific challenges arising in designing and implementing MP-based and MC-based schemes, such as, representation of realistic user characteristics (e.g., utility function, user priorities and cooperation), availability of information on users and traffic conditions, uncertainty in system conditions

and user behaviors, and circulation of mobility rights in market place. For the MP-based scheme, we focus specifically on designing a mobility scheme for single-bottleneck roadways. Roads with bridges, tunnels and business districts with limited parking spaces are the most obvious examples of a simple roadway with a single-bottleneck in a transportation network. We deal with observing operational objectives, specifically, balancing efficiency, equity (users priorities), and revenue outcome of distributing mobility permits under the "fairness" constraint. We explore the theoretical properties of the proposed scheme and show that the proposed scheme can achieve an optimal traffic pattern. Particularly, we show that the proposed scheme is a Pareto-improving and strategy-proof scheme capable of achieving efficient and effective market prices suitable for travelers. Our computational results indicate the effectiveness of the proposed scheme as an alternative solution for MP-based traffic management on single-bottleneck roadways.

We then investigate the case of traffic congestion management in a general road network through a MC-based scheme. Specifically, we propose a MC-based traffic management scheme in a road network consisting of a mixed-fleet traffic with connected and autonomous vehicles (CAVs) and conventional vehicles (non-CAVs). The basic premise of the proposed scheme is to regulate or influence travel demand and congestion with regards to the supply (capacity) of road networks, implementing a market-driven traffic management paradigm. A set of revenue-neutral, Pareto-improving MC-based charge and reward policies applicable to stochastic traffic environments are developed, considering different characteristics of users such as cooperative versus selfish routing behaviors, human-associated factors (e.g., level of uncertainty) and interactions due to a shared infrastructure setting. Path-free mathematical programming models are formulated, obviating computationally intractable path enumeration process pertinent to the existing studies. This makes the proposed scheme suitable for examining the theoretical characteristics of large-scale realistic transport networks. We examine several theoretical properties related to the proposed MC-based scheme, including the existence and uniqueness of the equilibrium price, and existence of Pareto-improving credit charges and rewards rates that can promote travel decision behaviors of individual travelers towards a network-wide optimal state. Our comprehensive computational results indicate that the proposed MC-based scheme can be an effective tool for managing travel demand and routing decisions in mixed-vehicle traffic settings.

Acknowledgments

I would like to express my deepest gratitude to my thesis supervisor Dr. Liping Fu for all his continuous support and his unbounded energy for being always ready to help and encourage me during these demanding years to make this thesis possible. I am thankful to you because of your trust in me; I am grateful for this amazing experience, providing me with this life-changing opportunity to join your research team. I believe, I was super lucky to get the chance to join and collaborate with a remarkable group of researchers in iTSS LAB and Railway Research Center (RRC) at the University of Waterloo without whom this thesis would not have been completed.

I would like to thank Professor Chris Bachmann, Professor Tarek Hegazi, Professor Jun Liu from the University of Waterloo and Professor Amy Miyoung Kim from the University of Alberta for agreeing to serve as examining committee member and for their time and effort in evaluating this thesis and for the challenging questions and providing me with valuable insights that improved the final exposition of my dissertation.

I would like to thank and appreciate my fellow colleagues and my friends at the University of Waterloo for their friendship and support. I would also like to thank the administrative staff of the Department of Civil & Environmental Engineering (CEE) and all other members of the University of Waterloo who have help me to fulfill this job.

Last but not least, I am extremely grateful to my parents for their unconditional love and support and my beloved family members who have been there for me always in my life and supported me unconditionally to get to where I am today. Words cannot describe how thankful to you I am.

Thank you all

To my mother, my father, and my beloved family members & To Ali and Jalal, and those innocent people of Ukraine Flight 752

Table of Contents

List of Tables

List of Figures

List of Acronyms

ATIS	*Advanced Traveler Information System*
BPR	*Bureau for Public Roads*
CAC	*Command-and-Control*
CAT	*Cap-and-Trade*
CAVs	*Connected Autonomous Vehicle*
CS	*Coordinated System*
DAT	*Desired Arrival Time*
DDT	*Desired Departure Time*
FCFS	*First-Come First-Served*
IC	*Incentive Compatibility*
IR	*Individual Rationality*
IoT	*Internet of Things*
KKT	*Karush-Kuhn-Tucker*
LP	*Linear Programming*
LPR	*License Plate Rationing*
MaaS	*Mobility-as-a-Service*
MC	*Mobility Credits*
MIP	*Mixed-Integer Programming*
MP	*Mobility Permits*
MSSO	*Mixed Stochastic System Optimal*
MSUE	*Mixed Stochastic User Equilibrium*
NCPs	*Non-linear Complementarity Problems*

OD	$Origin\ Destination$
PO	$Pareto\ Optimal$
PoE	$Price\ of\ Equity$
PoF	$Price\ of\ Fairness$
$ROWs$	$Right\ of\ Ways$
RP	$Random\ Priority$
SO	$System\ Optimal$
SSO	$Stochastic\ System\ Optimal$
SUE	$Stochastic\ User\ Equilibrium$
TAP	$Traffic\ Assignment\ Problem$
TTT	$Total\ Travel\ Time$
TTC	$Total\ Travel\ Cost$
$UMPAP$	$User\text{-}centric\ Mobility\ Permit\ Allocation\ Problem$
UE	$User\ Equilibrium$
VOT	$Value\ of\ Time$

Chapter 1

Introduction

1.1 Background

In roadway mobility networks, typically, each user individually looks for an ideal departure/arrival time and chooses a route and transportation mode that conceivably has the highest utility, under the prevailing traffic conditions. Without a proper mobility management strategy, individuals' self-concerned actions, in large, can lead to a collectively undesirable traffic state coined as congestion. It is one of the main factors contributing to inefficiency in transport with a broad range of externalities such as lost time, excessive wasted energy, social distress, public discomfort, harmful emissions and environmental deterioration.

In Canada, traffic congestion has reached an acute level. A conservative analysis conducted in 2006 showed that traffic congestion cost $4.6 billion annually, of which almost $3.7 billion was incurred in the Toronto, Montreal and Vancouver regions (Force 2012). A recent report by Commission et al. (2015) shows that congestion directly costs Toronto about $7 billion annually and Vancouver $1.4 billion. It is projected that the average annual cost to a commuter will grow to $1,100 and the national total congestion cost will grow up to $199 billion by 2020 (Schrank et al. 2012, 2015). Traffic congestion is also a common issue in many other countries around the world. For example, in urban areas of the US, each commuter approximately wasted $960 per year during the rush-hours on roads and the total congestion cost approximately $160 billion in 2015 (Schrank et al. 2015). A recent report shows that the average *travel time index* in the 10 most congested cities in China is close to 2, indicating that during rush-hours one would spend at least

twice the free flow travel time on a trip on a normal day. For Beijing, it was reported that the congestion cost totaled over \$11.3 billion in 2014, 80% of which was the lost time (Nie 2017a).

To harness the ever-increasing mobility demand in congested areas, practitioners and researchers have introduced a host of solutions, which can be classified into supply- and demand-based solutions. Supply-based solutions focus on increasing the capacity of transportation network by building roads and lanes and removing bottlenecks. However, increasing network capacity by adding more roads and lanes, turned out to be "self-defeating" in the sense that the increased capacity of road infrastructure can lead to increased travel demand and therefore congestion consequently (Johnston et al. 1995). Indeed, it has been shown that expanding road capacity is more likely to induce additional demand without reducing traffic congestion— the Pigou-Knight-Downs paradox (Downs 1962).

In contrast, demand-based strategies focus on reactive or proactive controlling of the travel demand from users side. There are two types of demand-based methods, namely, price-based and quantity-based schemes. The London congestion charge, the Stockholm cordon charge, and Singapore's Electronic Road pricing system are a few successful cases of congestion pricing. However, congestion pricing faces many questions pertaining to social equity and perception of fairness, which strongly influence the public acceptability and economic efficiency of the scheme (Chen and Yang 2012).

Though a powerful mechanism, congestion pricing has not been widely adopted due to public objections. Thus some researchers and planners have proposed quantity-based demand schemes that directly restrict mobility demand levels. Quantity-based regulations can be categorized into cap-and-trade (CAT) and command-and-control (CAC) schemes. CAT schemes are market-based regulatory mechanisms that assign right of ways (ROWs) along with economic incentives and tradability option while CAC set strict bounds on ROWs without integrating a fully free market dynamics such as tradability. Some variations of these types of schemes have been practiced in several cities across the world to control traffic congestion and related negative effects. The pre-established rules and car prohibition programs, such as alternate-day travel right, even-odd rationing, and generally license plate rationing (LPR), are indeed the simplest quantity-based CAC schemes. These control mechanisms have been applied in many Latin American cities like Mexico City and Sao Paulo, and recently in a few large cities in China such as Beijing and Guangzhou (Wang, Gao, Xu and Sun 2014b). These schemes, however, could become less effective in a long term. For instance, in Beijing and Mexico City it has been observed that people started

circumventing these rationing rules by owning multiple cars with odd and even numbers. This resulted in considerable welfare losses and inequality issues such as wealthier people can afford to own multiple cars. Consequently, Beijing abandoned the odd-even rationing policy, and introduced one weekday driving ban (Han et al. 2010, Gu et al. 2017). However, there has been no evidence that the "One Day without a Car" program has improved air quality or reduced the congestion significantly in Beijing (Davis 2008, Sun et al. 2014). In Bogota, capital of Columbia, cars were banned during the peak hours for two days per week in order to make it harder for citizens to break the rule by buying two cars. The government continued to switch the combination of days and numbers every year; however, people started to drive more during off-peak hours, thus rendering the government appointed peak hours as ineffective (Nie 2017a). Regardless of their configuration, the LPR schemes are proposed to reduce congestion by restricting the mobility demand via rationing, which can help reduce congestion under certain conditions. However, these policies are ineffective solutions as they lack supply-demand drivers and free market dynamics (Nie 2017b).

Mobility permit (MP)-based traffic management has been proposed as an alternative solution with potentials to circumvent the shortcomings of road pricing and simplistic rationing approaches. Mobility permit, also called mobility credit, is an innovative scheme that assigns roadway space to travelers according to pre-established permit endowment policies. The underlying idea is that one should have a permit or enough credits to use a specific road at a specific time, i.e., roadway space is a commodity or service to be bought from a transportation market. Some MP schemes allow the permit holders to trade or exchange their permits as "market goods" in a free market, which are thus coined as tradable mobility permit. This helps re-distribute social welfare among the eligible participants (Fan and Jiang 2013). It is worth mentioning that some of the aforementioned demand-oriented control schemes (e.g., LPR, even-odd and alternate-day rationing, and managed lanes) are actually primitive forms of permit-oriented mechanisms without integrating proper market drivers and users concerns within their operations. One of the main reasons for the limited adoption of MP-based schemes was lack of technological support and operational environment. This is however no longer the case due to recent technological advances and developments in the transportation sector as discussed in the following section.

1.2 Motivation

Mobility permit- or credit-based traffic management is a market-based mechanism to mitigate traffic congestion. Market-based schemes have been proven to be very effective solutions to similar problems in some other applications. For example, permit-based schemes have been applied to deal with public issues on common-property assets such as controlling pollution and emission and managing natural resources. A market-based traffic management mechanism can result in an increased net benefit for society by managing mobility demand in a collective way through which commuters are required to pay reasonable prices for transportation (Grant-Muller and Xu 2014). In other words, with a proper market-driven mechanism, the mobility demand and thus congestion can be controlled under the power of a free market medium. A road transport market can be constructed by casting road space as a common commodity; thus mobility demand can be controlled either with CAT or CAC schemes. The underlying idea is to endow the spatiotemporal ROWs to eligible commuters based on normative or rationing rules while observing the efficiency of the mobility system (Fan and Jiang 2013). Compared to the other traffic management strategies (e.g., tolling, pricing, or taxation) a permit-based mobility scheme is considered to be more flexible, targeted, cost-effective, and responsive to the changes in roadway transportation dynamics (Godard 2001). A properly designed and distributed permit-based scheme can attain a better social acceptability and allow for achieving a more desirable welfare state.

While the potential of these schemes has been shown and various permit- and credit-based management strategies have been proposed theoretically, none of them has been implemented and validated in practice due to the lack of advanced technologies and social and political support (Verhoef et al. 1997, Fan and Jiang 2013). This is no longer the case. Recent technological advances in information and communications technology (ICT) and automotive industry have paved the way for implementation of coordinative mobility permit-based traffic management system. Individual travelers are increasingly equipped with smartphones to communicate and exchange travel information. For example, many travelers are already using real-time mobile navigation applications, such as Google Maps[r] and Waze[r], to adjust their everyday mobility behaviors and make decisions about travel, route, mode of travel, and origin and destination choices. The introduction of connected and autonomous vehicles (CAVs) presents another opportunity for coordinated and cooperative travel, such as those in a MP-based mobility system (Klein and Ben-Elia 2016).

Another seismic shift that is taking place in the transportation sector is the rise of mobility marketplace and mobility-as-a-service (MaaS). MaaS aims to unify different travel options from different kinds and modes of mobility service providers, into a single seamless intuitive gateway such as a mobile App. It allows users to buy on-demand or subscribe to an affordable mobility package, and help them manage their trips from planning step to final payments (*Mobility-as-a-Service* n.d.). MaaS will bring about a huge change in the way people and goods move. People are expected to depart from personally-owned modes of transportation, and change their behavior from making conventional mobility decisions, from ride to here or there, to utilizing point to point mobility-on-demand services with a range of personalized mobility options. Mobility becomes an on-demand service requested by travelers who manage and use their credits to pay for consuming the limited capacity of the mobility network of all modes. This shift is expected to induce significant changes in the current transportation theories and practices. Expectedly, MaaS will be hosted by mobility service providers on top of smart infrastructures, mobility-sharing facilities, connected traffic systems, connected and autonomous vehicles (CAVs), tracking technologies along with Internet of things (IoT) through a unified gateway that serves travelers and manages their trips.

All these developments will have revolutionary impact on car ownership and trip making models, mobility demand, users' expectations about quality of service, and competition in the mobility market. Undoubtedly, with all the technological advances and the rising MaaS, MP-based traffic management system will become a reality. To enable such a future however, sophisticated service recommendation, permit allocations and pricing schemes must be developed, which are the main interest of this research, as detailed in the following sections.

1.3 Problem Statement

As discussed in the previous section, MP-based mobility management becomes increasingly feasible both technologically and politically. However, to assure the viability, effectiveness and acceptability of a permit-based traffic management system, many technical problems still need to be addressed. One of the main problems is how to price the mobility permits and allocate them to the users of different mobility systems. The underlying problem can be seen as a resource management problem, in which a decision maker needs to decide on the pricing and allocation of scarce resources (mobility permits as roadway usage rights)

to users considering different requirements and criteria. A pricing and allocation scheme must consider the objectives of mobility service planner and provider, mobility users' concerns and desires, and then determine the market values (prices) of permits according to transportation network characteristics (e.g., bottlenecks). Mobility permit pricing and allocation must balance between heterogeneous users' concerns about the fairness and mobility service providers' requirements about the overall efficiency of the system. In other words, given transportation network characteristics, a mobility manager must issue permits and endow them to the users of the system based on their time-varying desires while satisfying efficiency measures in terms of level of service and return on investment. In addition, the simplicity and computational efficiency of the scheme are other implementation-related factors that can highly affect the interoperability of the permit-based mobility management scheme.

The idea of using permits for mobility control can be traced back to Verhoef et al. (1997), who explored possible applications of different permit mechanisms in the regulation of road transport externalities. Goddard (1997) was the first to propose the use of mobility permits for restricting travel demand. Later, Koolstra (1999) studied the potential benefits of an advance slot reservation system for highway users, analyzing the difference between the user equilibrium and the system optimal departure times. Yang and Wang (2011) formally set up a mathematical model for MP-based traffic management problem in general networks with a link-specific charging scheme. They analyzed and explored a system of travel credits and obtained the existence of a unique equilibrium link flow pattern, with either fixed or elastic demand, by solving a standard traffic equilibrium model subject to a total credit consumption constraint. They also showed that, under a revenue-neutral manner, at equilibrium the credit price in the trading market is conditionally unique, and the correct selection of link-specific rates and appropriate distribution of credits among travelers can lead to the most desirable network flow patterns. However, their scheme is essentially a price-based mechanism that cannot eliminate congestion unless credit charges are time-varying or all travelers treat credit charge equivalently (Bao et al. 2016). Moreover, it is limited in addressing users' priorities and concerns on the equity and fairness notion. In addition, their scheme faces the asymmetric information problem which can lead to a regulatory failure. Doan et al. (2011) addressed the pricing strategies in the discrete time single bottleneck model with general heterogeneous commuters and proved that a system optimal assignment can eliminate the queue. They formulated the system optimal problem as a linear programming (LP) model and discussed the existence and uniqueness of solution by applying duality theory. They proved that an equilibrium solution can

be achieved with corresponding optimal tolling. Extending the work of Yang and Wang (2011), Wada and Akamatsu (2013) proposed a network scale MP-based mechanism under which each commuter purchases a specific number of permits based on their preferred path. However, their mechanism requires a commuter to buy all the link-specific permits along their path which makes it very complicated for a real-world transportation network. All these studies also assume that users behave rationally and react consistently under the same assumed settings, which is not the case in practice.

To address travelers' self-concerned behavior, some of the proposed MP-based schemes try to operate the transportation network at a user equilibrium (UE) traffic pattern where travel decisions are coordinated through a unified gateway. On the other hand, to address the mobility service providers' concerns about efficiency of the system some researchers proposed using MP-based schemes that operate the mobility system at a system optimum state. However, neither UE is the most efficient state when the collective travel times in the entire transport system is expected to be minimized, nor system optimal (SO) is a stable equilibrium in the absence of proper intervening forces such as incentives and charges (Rothengatter 1982). To manage mobility demand efficiently in a transportation network, it is important to account for the dynamics of users' behavior and expectations on top of the simplicity and efficiency concerns. In our understanding about MP-based scheme, more attention is being given to UE and SO measures which should not always be the sole or major considerations. In this regard, a coordinated mobility control based on a user-centric permit scheme can gain more political and public support at a satisfactory level of network efficiency. The proposed designs are also limited in addressing application-related factors which have hindered the practicality a MP-based schemes. For instance, most of the current studies are built on restrictive assumptions, e.g., rationality of travelers and availability of perfect information. They also assume that in a transportation network, individual users choose a path that conceivably has the least cost (time) from the origin to the destination which is not the case always in practice. A mechanism that relaxes these assumptions seems to be more requested in practice to facilitate the adoption of permit-based mobility management systems in large.

A recent survey in the San Francisco Bay Area revealed that about 60% of the respondents were willing to consider user-specific mobility services (Khattak and Yim 2004). This means that on our road to a future with a mobility ecosystem, we need to be equipped with innovative mobility control policies while putting users at the core of transport services. Therefore, we need to incorporate issues and concerns both from travel demand and mobility supply sides to make a mobility permit-based traffic management scheme operable. We

must help users feel secure and trust a permit-based mobility system and satisfy operators and investors expectations on return and efficiency. The prospect should be developing futuristic mobility management measures, aiming to improve efficiency and effectiveness of current and future mobility systems. These forms of mobility service mechanisms should be of interest to both users and operators.

1.4 Research Objectives

This research is motivated by the need to develop innovative mobility management systems for mitigating urban traffic congestion in transportation networks with advanced connectivity technologies. In particular, we will consider a set of futuristic mobility management schemes to control travel demand and traffic congestion in the emerging transport technologies. The proposed research will focus on some of the main technical issues with MP-based, such as trade-off between efficiency and fairness, users' preferences, uncertainty and heterogeneity, and computational efficiency needs. We will also look into mixed-autonomy transportation networks with cooperative and non-cooperative travel behavior of users. Specifically, the objectives of this research are to:

- Construct an integrated framework considering issues related to system settings and requirements of a futuristic user-centric mobility management scheme,

- Develop mobility management schemes for single-bottleneck roadways and network with multiple bottlenecks,

- Develop models and solution algorithms for the underlying pricing and allocation problems of the proposed mobility management schemes with an explicit consideration of equity and efficiency requirements,

- Explore the theoretical properties of the formulated problems, such as existence of equilibrium and optimality, and investigate the effect of various assumptions about the system settings,

- Conduct case studies to demonstrate the applicability and effectiveness of the proposed models and algorithms with comprehensive experiments on system performance and sensitivity under different settings.

1.5 Thesis Outline

This thesis is comprised of six chapters as follows:

- Chapter 1 introduces the motivation and background of the proposed research, and continues with highlighting current practices, directions in the mobility management field and limitations of current studies.

- Chapter 2 provides a literature review, classification and taxonomy of various mobility management schemes proposed in the relevant studies.

- Chapter 3 discusses the proposed conceptual framework of an integrated paradigm on designing mobility permit-based travel demand management scheme to highlight the operational challenges of designing futuristic user-centric mobility management schemes.

- Chapter 4 focuses on the design and performance analysis of mobility permit pricing and allocation on single-bottleneck roadway. It presents the theoretical properties of the proposed scheme and provides the results of a comprehensive set of numerical experimentations under different settings.

- Chapter 5 focuses on the design and performance analysis of a credit-based mobility management scheme for transportation networks with ATIS and connectivity technologies. It presents the theoretical properties of the proposed scheme and provides the results of comprehensive numerical experimentations under different settings.

- Chapter 6 concludes the study, summarizes the findings of the research discussed in chapters 3 to 5 and draws conclusions based on this work. Important next research steps are highlighted as well.

Chapter 2

Literature Review †

2.1 Summary

With the advent of new communication and information technologies, the idea of treating roadway transportation as a shared economy and next-generation of mobility services is increasingly becoming a reality. The basic premise is that the roadway system can be viewed as a collection of scarce commodities of a shared economy and it can be managed by endowing right use to its users. Over the past two decades, a growing body of research has explored different variations of roadway-use right schemes, including credit- and permit-based mobility schemes, with specific focus on the three problems arising in these schemes—pricing, efficient allocation and charging of permits or credits for mobility management. In this chapter, we attempt to synthesize the relevant body of literature by presenting an in-depth comprehensive review of the state-of-the-art methodologies for addressing these decision problems. The goal is to provide an improved understanding of the fundamental constructs of these problems by systematically classifying the problem variants, proposed methodologies, and modeling attributes. We also discuss the research gaps and challenges and suggest some possible perspectives and directions for future research.

This chapter surveys the existing contributions in the literature on the methodological advancements on permit- and credit-based mobility management schemes, focusing on the

---

†This chapter is adapted from a previously published paper: Lessan, J., & Fu, L. (2019). Credit-and permit-based travel demand management state-of-the-art methodological advances. Transportmetrica A: Transport Science, (just-accepted), 1-24. by Taylor & Francis in Transportmetrica A: Transport Science on 2019/11/14, available online: *https://www.tandfonline.com/doi/full/10.1080/23249935.2019.1692963.*

use of different models and problem settings. Great attention is also given to the identification of the potential research directions for future researchers. The specific objectives of this chapter are to:

1. review the contemporary methodological advancements on congestion management using roadway-use right schemes;

2. provide an improved understanding of the fundamental constructs of the specified problems;

3. synthesize the relevant body of literature and classify different proposed methodological approaches; and

4. discuss the research gaps, challenges, and possible directions for future research.

2.2 Background

The idea of rationing a limited common resource (capacity) through permits can be traced back to the theory of permit markets in regulating the use of the common resources, which was suggested by Crocker (1966) and Dales (1968) as a way of controlling atmospheric pollution and water pollution respectively. In roadway transportation sector, some early studies on roadway-use rights mainly focus on the conceptual developments, potential benefits, and feasibility of different credit- and permit-based control schemes for mitigating road transport externalities (Goddard 1997, Verhoef et al. 1997, Kockelman and Kalmanje 2005, Gulipalli and Kockelman 2008). However, none of these studies address the key decision issues arising in the design and implementation of these schemes—pricing, distributing, and charging problems— and their influences on traffic flow and market equilibrium conditions. Since then, a resurgent interest has focused on developing different variations of roadway-use right schemes and related issues.

2.3 Problem Description and Formulation

As discussed previously, roadway-use right schemes can be classified into two distinct approaches, namely, credit- and permit-based. A permit-based scheme involves a road authority restricting roads (or links) usage to those travelers who have acquired road-specific

permits in advance (Wang, Liu and Huang 2018). In credit-based schemes, on the other hand, the road authority would issue a certain number of mobility credits to all eligible travelers and then charges the users with a specific number of credits for each road (link) they have used (Yang and Wang 2011). A permit is issued to a specific user for a specific time interval or place, while credits are user agnostic and can be used for any time intervals or places. Another difference is that a credit-based scheme does not directly restrict travel demand, while as a quantity-based mechanism, a permit-based scheme often aims to completely eliminate the occurrence of traffic congestion by not letting the travel demand at any time interval to be greater than the capacity of the network (Wada and Akamatsu 2013).

2.3.1 Problem Context

The roadway-use right schemes generally involve three different actors (stakeholders): transportation firms (e.g., a private service provider), road network (mobility) users, and central transport authority (government). The transport authority aims to mitigate traffic congestion on the network and needs to deal with not only individual users but also transportation firms and transit agencies. In the mobility right system, each user must purchase a set of permits or must acquire enough credits corresponding to the chosen mobility service. The roadway-use rights are good instruments to commoditize mobility services and spread social welfare by letting the participants exchange their rights, allowing the transport authority to outsource mobility services to third party, i.e., transport firms. The private firms can also finance roadway construction and receive a part of the equities as a reward for constructing the road network system (Yang and Wang 2011), or they can be transport operators offering a broad range of mobility services. The central transport authority needs to determine permits or credits distribution to the users, and subsequent usage charges. A roadway-use right scheme from the transport authority's perspective can have different goals, such as minimizing total travel time (TTT) or total transportation cost (TTC), or maximizing the social welfare of a transportation system. All these require the transport authority to understand the key decision problems to manage the travel demand in the network so as to achieve some system-level goals.

2.3.2 Key Decision Problems

Despite all the differences, the implementation of the credit- and permit-based schemes faces similar decisions: how should the credit/permit be distributed, priced, and charged?

The first problem is of resource allocation, in which the central transport authority decides on the allocation or distribution of scarce resources (usage rights) to users, including how many permits or credits to issue and to whom. This task, however, depends on several factors such as OD demands, network characteristics, bottlenecks, links dependencies, and time periods. The second problem is how to price the mobility rights which is also affected by users' travel behavior, such as arrival/departure time, mode, route choice and trading behavior. The third problem is the charging mechanism for the usage of the roads (links), which depends on the scheme itself and the network characteristics.

2.3.3 Basic Formulations

In what follows, we introduce the basic formulation of roadway-use right schemes, focus on their similarities and differences. This will provide a foundation for discussing various extension of the basic models in the subsequent sections.

Credit-based Mobility Scheme

To formally illustrate the traffic flow pattern and network equilibrium conditions under this scheme, we borrow the credit-based model of Yang and Wang (2011) for a roadway transportation network $G = (N, L)$ with a set of N nodes and a set of L directed links. In this network, each node is identified by a sequence of natural numbers i, and each link of L is denoted by a pair (i, j) of the upstream node i and the downstream node j. Let the set of OD pairs be denoted by W and R_w be the set of all routes between an OD pair $w \in W$. The travel demand for each OD pair $w \in W$ is denoted by q_w, $q_w > 0$. The transportation authority implements a network-wide credit-based travel demand management system in which a specific number of credits are distributed to individual road users in advance and each user is charged with a link-specific amount of credits for using each link. The road users are assumed to be homogeneous, each having a value of time (VOT) equal to unity.

It is assumed that the total number of available credits are distributed uniformly to each traveler over the OD pair $w \in W$; that is, the number of credits available to all

users of the same OD pair is the same. Let K denote the total amount of credits, and $\bar{k}$ denote the initial credit amount distributed to each traveler, thus $K = \bar{k}\sum_{w \in W} q_w$. Let k_{ij} denote the credit charges on link $(i, j) \in L$; thus $k = k_{ij}, (i, j) \in L$ denotes a link credit charge scheme. Therefore, the entire credit charging scheme can be represented by (K, k). With this credit-based scheme in place, all users would choose their routes based on their generalized travel cost (τ_w), which is the sum of total travel time plus the link-dependent credit charges, in such a way that their generalized travel cost for their trips are minimized while they have sufficient credits for traversing the links along their chosen routes. Indeed, at traffic equilibrium, all utilized paths between the same OD pair $w \in W$ will have equal and minimal generalized travel cost and for all the utilized paths the generalized travel cost should be greater than the minimal generalized travel cost. At the same time, the credit market price (p) is positive only when all the issued credits are consumed (Yang and Wang 2011). The resulting state can be formulated as mathematical model (2.1)-(2.3), integrating UE and credit market equilibrium state.

$$\min_{\boldsymbol{v} \in \varphi_{(\boldsymbol{f},\boldsymbol{v})}} Z(\boldsymbol{v}) = \sum_{(i,j) \in A} \int_0^{v_{ij}} t_{ij}(\omega)d\omega \tag{2.1}$$

$$subject\ to:$$

$$\sum_{(i,j) \in A} k_{ij} v_{ij} \leq K, \tag{2.2}$$

where $\varphi_{(\boldsymbol{f},\boldsymbol{v})}$ is the feasible set of OD demand and link flows, which can be defined by:

$$\varphi_{(\boldsymbol{f},\boldsymbol{v})} = \{(\boldsymbol{f}, \boldsymbol{v})|, v_{ij} = \sum_{w \in W} \sum_{r \in R_w} f_r^w \delta_{ij,r}^w, \sum_{r \in R_w} f_r^w = q_w,$$

$$f_r^w \geq 0, \forall (i, j) \in L, \forall w \in W, \forall r \in R_w\}, \tag{2.3}$$

where f_r^w denotes the traffic flow on path $r \in R_w$, v_{ij} denotes the traffic flow on link $(i, j) \in L$, and $\delta_{ij,r}^w$ is equal to 1 if path $r \in R_w$, $w \in W$ uses link $(i, j) \in L$ and zero otherwise. For simplicity, let $t_{ij}(v_{ij})$, $\forall (i, j) \in L$ denote a separable and monotonically increasing link travel time function; therefore, the path travel time on route $r \in R_w$ can be obtained as $t_r^w = \sum_{(i,j) \in L} t_{ij}(v_{ij})\delta_{ij,r}^w, r \in R_w, w \in W$. Assume that $\varphi_{(\boldsymbol{f},\boldsymbol{v})}$ is non-empty, according to Yang and Wang (2011) for this convex optimization problem with linear constraints an optimum solution $(\boldsymbol{f}^*, \boldsymbol{v}^*)$ is obtainable with the Lagrange multipliers p and $\boldsymbol{\tau} = (\tau_w, w \in W)$ that are associated with the credit feasibility and the path flow

conservation constraints, respectively. The Lagrangian function L for the problem can be expressed as

$$L(\boldsymbol{v}, \boldsymbol{\tau}, p) = Z(\boldsymbol{v}) + \sum_{w \in W} \tau_w \{ \sum_{r \in R_w} f_r^w - q_w \} + p(K - \sum_{(i,j) \in A} k_{ij} v_{ij}). \qquad (2.4)$$

The general optimality conditions for the problem can be obtained by the Karush-Kuhn-Tucker (KKT) conditions. Indeed, the optimality conditions are equivalent to the UE conditions, and the credit market equilibrium conditions (Yang and Wang 2011). Moreover, the Lagrangian multipliers p and τ_w correspond to the unit credit price at market equilibrium and the minimal generalized travel cost at equilibrium respectively. Nonetheless, as discussed there are some special cases where the credit price may not be unique. This problem can be extended to investigate market equilibrium and traffic flow patterns with credit-based scheme under social optimum, Pareto-improving (making each individual user better off) and revenue-neutral (having equal charges and subsides), and restrained traffic flow situations (Yang and Wang 2011).

Permit-based Mobility Scheme

We borrow the model of Wada and Akamatsu (2013) and Akamatsu and Wada (2017), to formally illustrate the permit-based mobility management scheme. The general network consists of a set of N nodes, and a set of L directed links. Each element of N (i.e., each node) is identified by a number i, and each element of L (i.e., each link) is denoted by a pair (i, j) of the upstream node i and the downstream node j. For simplicity, it is assumed that the network contains only a single OD pair in which R is the set of path connecting the OD pair (o, d). The total travel demand, Q, is fixed during the time interval $[0, T]$. Let $q(t)$, $q(t) > 0$, denote the OD flow for users arriving at the destination at time period t.

To control the travel demand, the transportation authority implements a link-based mobility permit system in which each user must hold a permit in order to use a specific road (link) in the network. Without loss of generality, it is assumed that the number of permits issued by the authority for each link (i, j) over each time period is limited so that the total traffic (demand) using the link would not exceed its capacity (μ_{ij}) at all times. Moreover, it is assumed that for each link (i, j) the unit price of the permits of link (i, j) is a function of the time the user enters the link t, denoted by $p_{ij}(t)$. Therefore, this scheme

would lead to varied permit prices in the market. With this permit-based scheme in place, all users would choose their routes and travel times in such a way that their generalized transportation cost for their trips are minimized while the conservation of dynamic link flows and bottlenecks (capacities) conditions are sustained. The generalized transportation cost consists of the schedule cost which is represented by $w(t)$ as the function of destination arrival time period t and the monetary equivalent of the total travel time paid by all users. Indeed, at traffic equilibrium no user can improve his/her own cost by changing the path choice unilaterally, i.e., all utilized paths between the same OD pair $w \in W$ will have equal and minimal generalized travel cost and for all the unused paths the generalized travel cost would be greater that the minimal generalized travel cost. At the same time, the permit market price of each link (p_{ij}) is positive only when the link is saturated. The equilibrium state can be obtained from the solution of a mathematical formulation similar to the credit-based scheme. However, due to the relative complexity of the permit-based scheme, we separately illustrate each of its component and then show how they together build up to the final model.

To describe travelers' choices and the conditions on link flows and capacity conservation, let $y_{ij}(t)$ $(z_{ij}(t))$ be the inflow (outflow) arriving (leaving) at (from) link (i, j) at time period t. Moreover, let $NO(i)$ $(NI(i))$ denote the set of downward (upward) nodes of the links incident to (from) node i, then the flow conservation for each node i can be represented as

$$\sum_{k \in NO(i)} y_{ik}(t) - \sum_{k \in NI(i)} z_{ki}(t) = -q(t)\delta_{id}, \ \ \forall t \in I, \forall i \in N, \tag{2.5}$$

where $\delta_{id} = 1$ if $i = d$; zero otherwise. Without loss of generality, it is assumed that the time interval $[0, T]$ is discretised into small intervals $[t, t + \Delta t]$ of length Δt. Each time interval is represented by $t = m\Delta t$, where $m = 0, 1, 2, ..., M$ and denoted as interval time period t for $t \in I$.

Furthermore, it is assumed that the dynamic traffic flow on each link should satisfy the First-In-First-Out (FIFO) condition which is expressed as

$$A_{ij}(t) = D_{ij}(t + t_{ij}(t)), \tag{2.6}$$

where $A_{ij}(t)$ $(D_{ij}(t))$ is the cumulative numbers of permit holders arriving (leaving) into (from) link (i, j) at time t. Using the flow variables, the FIFO condition can be written as

$$y_{ij}(t) = z_{ij}(t + t_{ij}(t)).(1 + dt_{ij}(t)/dt), \tag{2.7}$$

where $t_{ij}(t)$ is the travel time of link (i,j) for a user entering into the link at time t. Given that $t_{ij}(t)$ is a constant under the permit-based scheme where there is no congestion in the network, the FIFO condition reduces to

$$y_{ij}(t) = z_{ij}(t + t_{ij}(t)), \ \ \forall t \in I, \forall (i,j) \in L. \tag{2.8}$$

The conservation of dynamic link flows condition (2.5) at each node combined with the FIFO condition (2.8) on each link can be expressed by

$$\sum_{k \in NO(i)} y_{ik}(t) - \sum_{k \in NO(i)} y_{ki}(t - t_{ki}) = -q(t)\delta_{id}, \ \ \forall t \in I, \forall i \in N. \tag{2.9}$$

The bottleneck constraint on each link (i,j) with capacity limit μ_{ij} can be expressed as

$$y_{ij}(t) \le \mu_{ij}, \ \ \forall t \in I, \forall (i,j) \in L. \tag{2.10}$$

Finally, flow conservation for OD flows and OD travel demand can be written as

$$\sum_{t \in I} q(t) = Q. \tag{2.11}$$

Under this system of permit-based mobility scheme, the dynamic traffic flow pattern that minimizes the total generalized transportation cost on the feasible set of time-dependent OD demand and link flows conservations can be modeled as

$$\min_{(\boldsymbol{q}, \boldsymbol{y}) \ge 0} F(\boldsymbol{q}, \boldsymbol{y}) = \sum_{t \in I} q(t)\omega(t) + \alpha \sum_{(i,j) \in L} \sum_{t \in I} y_{ij}(t) t_{ij} \tag{2.12}$$
$$subject\ to:$$
$$(2.9), (2.10), and\ (2.11).$$

The first term in the objective function is the so-called total schedule cost, and the second term, with α being a coefficient that converts travel time to the monetary equivalent,

is the total travel cost. For any network in which the above problem has feasible solutions, Akamatsu and Wada (2017) show that the equilibrium assignment under the system of time-dependent link-specific permits minimizes the 'social transportation cost' defined by the objective function. This can be shown using the necessary and sufficient conditions for the optimality of the optimization problem. To derive the optimality conditions, the Lagrangian function L for the problem can be expressed as

$$
\begin{aligned}
L(\boldsymbol{q}, \boldsymbol{y}, \rho, \boldsymbol{\pi}, \boldsymbol{p}) \quad = \quad & F(\boldsymbol{q}, \boldsymbol{y}) + \\
& \sum_{i \in N} \sum_{t \in I} \pi_i(t) \{ q(t) \delta_{id} + \sum_{k \in NO(i)} y_{ik}(t) - \sum_{k \in NO(i)} y_{ki}(t - t_{ki}) \} + \\
& \sum_{(i,j) \in L} \sum_{t \in I} p_{ij}(t) \{ y_{ij}(t) - \mu_{ij} \} + \rho(Q - \sum_{t \in I} q(t)),
\end{aligned}
\tag{2.13}
$$

where function F is the objective function of the optimization problem; $\boldsymbol{\pi}, \boldsymbol{p}, \rho$ are Lagrangian multipliers corresponding to constraints (2.9)-(2.11) in the optimization problem, respectively. The necessary and sufficient conditions for the optimality of the problem given by the KKT conditions would result in the optimal values of $\boldsymbol{p}^*(t)$, $\boldsymbol{\pi}^*(t)$, ρ^*. The optimal value of these multipliers coincides with the equilibrium link permit prices, equilibrium minimum path costs, and the equilibrium generalized transportation costs in equilibrium conditions. The optimal flow patterns $(\boldsymbol{q}^*(t), \boldsymbol{y}^*(t))$ also coincide with the equilibrium flow patterns (Akamatsu and Wada 2017). The author also extend the proposed scheme to general networks with many-to-many OD pairs under different conditions such as heterogeneous users with different schedule delay functions and users with elastic trip demands cases.

2.4 Classification of Roadway-use Schemes

The basic problems described in the previous section have been extended in many directions over the past decades. Many variations of roadway-use schemes and problem formulation have been proposed. In this section, we first provide a classification of different variations and extensions of the roadway-use right schemes. We then look into different modeling characteristics and specific transport setting used in the literature. Table (2.1) and (2.1) summarize the state-of-the-art models on roadway-use right schemes given different attributes and assumptions. We have identified some differences and similarities in the at-

tributes used in the proposed schemes, including credits/permits distribution, charging and transferability, users' characteristics, system state, level of information, transaction costs, financing, sustainability, parking space regulation, and the effect of CAVs on roadway-use schemes. In the following subsections, we provide a detailed discussion about each of these specific issues.

2.4.1 Credits/Permits Distribution, Charging and Transferability

As discussed earlier, an important issue with the existing credit- and permit-based schemes is that they require the transport authority to predetermine the credits or permits charging levels. Recently, Wang, Gao and Xu (2019) proposed a mixed-integer non-linear bi-level programming model, which includes an upper level sub-model for the transport authority that tries to find the optimal number of lanes and credit charging level with their locations, given a lower level sub-model for user equilibrium formulation. In the sub-model the network users try to minimize their individual generalized travel cost given the network design strategy determined by the upper level problem. Another issue with the credit- and permit-based schemes is that they need well-designed periodic credit/permit distribution and collection strategies by the transport authority. To circumvent this issue, Xiao, Long, Li, Kou and Nie (2019), present a credit internal cycling without periodic credit expiration and distribution. They examine conditions for the existence of the solution, and prove that there is no negative cycle under the proposed scheme and that the Pareto-improving solution exists under certain conditions. Our review shows that almost all of the studies consider the tradability of the permit or credits in their scheme designs to turn the mobility rights into a market commodity, complementary incentive or equity improvement measure, except for Liu et al. (2015) that applies a reservation scheme for travel permit and Liu and Nie (2017) that allows users to redeem or rebate their unused credits. Regarding the distribution mechanism, the roadway-use schemes can differ in the way they distribute the credits or permits to users. While most of the proposed schemes bestow all eligible travelers with equal mobility charges or rewards, some studies propose different right endowment process where the mobility rights are tied with license plate that directly controls auto ownership (Nie 2017*b*). This confirms that giving mobility rights to all (eligible) travelers is possibly more efficient than the other alternatives.

2.4.2 Users' Characteristics

The most typical setting considered in literature assumes homogeneous commuters/vehicles with identical travel time or route duration. However, user heterogeneity is a central issue in traffic equilibrium analysis and plays a pivotal role in shaping the property and design of roadway-use schemes (Zhu et al. 2014). Doan et al. (2011) address the pricing strategies in the discrete time single bottleneck model with general heterogeneous commuters. Wang et al. (2012) extend the work of Yang and Wang (2011), considering heterogeneous users with different VOTs. They formulate it as a variational inequalities problem and examine the sufficient conditions for the uniqueness of the aggregate user equilibrium link flows and establishment of market equilibrium credit price. For managing bottleneck congestion and modal split in a competitive highway/transit network with continuous heterogeneity in the individuals' VOT who are initially endowed of a certain amount of travel credits, Tian et al. (2013) implement time-dependent credit charges only for the usage of the road bottleneck in a competitive highway/transit network with continuous heterogeneity in the individuals' VOT.

Zhu et al. (2014) investigate the multi-class network equilibrium problem under a credit scheme. They treat the heterogeneity of users by adopting a continuously distributed VOT in their evaluation of travelers' time savings. In the spirit of analyzing the effects of commuters' characteristics in determining link tolls and total allocated credits, Wang et al. (2012) consider heterogeneous users with a discrete set of values of time. They investigate the relationship between the uniqueness of the aggregate UE link flow pattern and the equilibrium credit price. He et al. (2013) study the effect of the mixed behaviors of UE-following and self-optimizing oligopoly Cournot players in the optimal design of a credit-based scheme with transaction costs. Bao et al. (2014) develop a more realistic scheme by considering travelers' loss aversion behaviors in their route choice. Given the market with transaction cost for buying and selling credits, they demonstrate that the system optimum link flow pattern may not be achievable when travelers' loss aversion behavior is considered. Zhu et al. (2014) investigate the conditions under which travelers are assumed to be heterogeneous with a continuous distribution of value of time. Nie and Yin (2013) propose a scheme to manage morning commute choices where the central authority divides the planning horizon into peak and off-peak periods, rewarding off-peak period commuters and charging peak period commuters. Liu et al. (2015) show that user heterogeneity causes further loss of efficiency, and thus they propose an auction-based reservation to mitigate the efficiency loss. Akamatsu and Wada (2017) propose an auctioning system for a designated

bottleneck where vehicles can use it with time-place specific permits.

Wang, Liu and Huang (2018) propose a OD-based travel permits scheme to manage mobility in bi-modal networks with user heterogeneity in VOT, which simplifies the static model setting to the homogeneous case. The modal split equilibrium is guaranteed by modelling the market-clear conditions as a complementary problem. The proposed scheme is actually a second-best pricing approach as all travellers between the same OD pairs are charged the same permit, independent of their respective used paths and links (Wang, Liu and Huang 2018).

Elasticity is another factor that affects the efficiency of the travel management policies. Yang and Wang (2011) show that under an elastic demand case, a unique user equilibrium flow pattern can be obtained too. Zhu et al. (2014) show that their modeling framework can be used to deal with demand elasticity by constructing an expanded network. Bao et al. (2016) look into traffic assignment under elastic demand case. The elastic demand case is also considered in a few other studies as well (Miralinaghi and Peeta 2016, Wang, Liu and Huang 2018, Bao et al. 2017). For the permit-based scheme, Akamatsu and Wada (2017) investigate the case with elastic demands where OD demand is a monotone decreasing function of the generalized transportation cost. Some studies look into another notion of elasticity, internal elasticity, by allowing inter-modal competition or substitution of modes with respect to credit charges (Xu and Grant-Muller 2016*b*,*a*, Tian et al. 2013). Specifically, Tian et al. (2013) investigate trip mode and travel pattern impacts on credit schemes and show that when the system optimum is achieved the total social cost is reduced for all transit modes.

2.4.3 System State

Most of the models in the literature are "single-period" static schemes. However, some studies have taken a multiple period setting into account. Here, we only look into multi-period credit- or permit-based mobility schemes. Ye and Yang (2013) propose a day-to-day dynamic model assuming that the credit price on each day is a function of the credit price of the previous day and the excess credit demand in the market. Miralinaghi and Peeta (2016) is the first study to address a long-term planning problem with the concept of a multi-period permit-based scheme. Wang, Wada, Akamatsu and Nagae (2018) extend the theory of bottleneck permits of Akamatsu and Wada (2017) to cases with multiple period markets, assuming that the users' valuations for permits depend on the purchase period.

They design its implementation mechanism and show that a multiple period market can be more efficient than a single period market when users' valuations change over time.

While the initial works on designing the credit- and permit-based schemes assume travel demand, users behavioral choices, and network condition to be fixed, many elements of an actual system could exhibit significant uncertainties. To ensure the designed schemes are robust against changes, some short-term uncertainties such as day-to-day time-varying travel demand and users choice behavior, long-term changes in land-use, and technological advances should be accounted in the schemes. Complexity is introduced when uncertainty is considered in roadway use schemes. Han and Cheng (2017) establish an equivalent minimization model for the stochastic user equilibrium assignment problem with a credit scheme. They consider the travelers' perception errors of generalized path travel costs where all the available routes are likely to be chosen according to a certain probability distribution.

The majority of articles studied in this survey consider static and deterministic settings and rely on the restricted assumption that model inputs such as travel demand and link capacities are time-invariant. However, daily volatility of traffic flow and fluctuation of network's links capacities can influence travelers' route choice behavior. For a continuous dynamic model in a finite time horizon the existence and uniqueness of the equilibrium are established in Ye and Yang (2013) who examine the price and flow dynamics under the travelers' learning of the evolution of network flows and credit price. Under a no late-arrival assumption, a dynamic credit charging scheme is proposed in Xiao et al. (2013) to build a similar scheme to that of Yang and Wang (2011) where the credit price is determined by a competitive market. Irrespective of how commuters vary in their value-of-time, the authors show that a time-varying credit charging scheme can eliminate traffic congestion at bottlenecks. A time-varying credit charging scheme is also proposed in Tian et al. (2013) and Tian and Chiu (2015) where users are able to form individual propensities in their travel decisions by reinforcement learning principles. Miralinaghi and Peeta (2016) focus on the fluctuation of credit prices over the planning horizon rather than within a period. The proposed model can be combined with day-to-day models of Ye and Yang (2013) to represent credit price and link flow evolutions at each period. Li, Ukkusuri and Fan (2018) cast light on how to schedule a time-varying credit charge scheme. An optimal dynamic credit scheme is proposed to attain mobility and emission goals by redistributing travel demand flows. To capture the flow propagation and dynamic user equilibrium a dynamic network loading model is proposed to investigate the flow redistribution in terms of simultaneous path and departure time choices under a carbon credit charge scheme.

2.4.4 Level of Information

Travel demand, link travel times and travelers' VOTs are not usually readily available in practice (Yang and Wang 2011). However, accurate estimation of potential travel demand and users' preferences can have a significant impact on the efficiency of a roadway scheme. Therefore, the regulator or service providers need relevant information about crucial factors such as users' VOT and travel demand (Grant-Muller and Xu 2014). Though it is proven that under symmetric information setting both Pigouvian taxes and credit schemes that allow for tradability yield the same production efficiency outcome (Bao et al. 2017). However, in real-world asymmetric information setting both can lead to regulatory failure and underpricing. One stream of research has addressed some of the uncertainties involved in the design of roadway-use right schemes. For the single roadway setting, Wang and Yang (2012) illustrate the effectiveness of a revealed credit price in determining the optimal credit scheme in the absence of a demand function and propose an iterative credit adjustment procedure with a guaranteed linear convergence rate. However, in a general network, the problem becomes much more complicated. Han and Cheng (2015b) investigate the effects of the travelers' perception errors on the travel cost under mobility schemes, which corresponds to the stochastic user equilibrium. They establish an equivalent minimization model with linear constraints for UE and market equilibrium conditions. It is then simplified into the UE model with a linear total credit amount constraint, under the assumption of Gumbel distributed perception errors. Then, the Wardrop system optimal mobility scheme problem is analyses, and a sufficient condition is provided for the system.

2.4.5 Transaction Costs

Roadway-use scheme could require significant administrative effort for managing credits or permits. To investigate the impact of administrative burden, Nie (2012) looks into the impacts of transaction costs on traffic and market equilibrium and the system efficiency of permit-based model of Yang and Wang (2011). He examines the effects of transaction costs on auction and negotiated markets. In an auction market, users purchase all of the needed mobility credits through competitive biddings. In a negotiated market, the users can trade the initially received amount of mobility credits from the government with each other through negotiation to meet their needs. He shows that with proper prices the auction market can achieve the desired equilibrium allocation of mobility credits as long as the unit transaction cost is lower than the price that the market would reach in absence of

transaction costs. He et al. (2013) study the effect of the mixed behaviors of UE-following and self-optimizing oligopoly Cournot players in the optimal design of a credit scheme with transaction costs. Bao et al. (2014) demonstrate that when travelers' loss aversion behavior is considered the system optimum link flow pattern may not be achievable for a market with transaction cost.

2.4.6 Financing and Sustainability

The involvement of transportation firms, as an integral stakeholder in the markets of travel, has been considered in Bao et al. (2017). They examine private financing and mobility management of road network with a build-equity-credit (BEC) scheme. The properties of several different BEC scenarios are investigated and it is found that the link service level in BEC is not constant but depends on multiple factors and the total market value of the credits charged on the new link can offset its construction cost and the profit of the private firm can always be non-negative. In the same spirit, Wang, Gao, Xu and Sun (2014a) investigate a public-private partnership network charging plan with a hybrid implementation of credit- and road pricing schemes. The charging plan comprises a credit scheme for public roads and a regular tolling sub-scheme that is imposed in the sub-network of private roads under build-operate-transfer contracts. Taking into account Cournot-Nash players, the authors develop three bi-objective optimization models with hybrid roadway charging schemes and show that there exist anonymous feasible hybrid charging schemes that support a system optimum link traffic pattern.

Wang, Xu, Grant-Muller and Gao (2018) combine a credit-based mobility scheme and link capacity improvement measure that orients sustainable transport development by considering road supply (i.e. link capacity) development and travel demand management. Similarly, in the recently published paper (Wang, Gao and Xu 2019), a link-specific credit charging scheme is integrated into the network design problem to better the transport performance from the both transport network planning and travel demand management perspectives. In a similar vein, Akamatsu and Wada (2017) and Akamatsu et al. (2006) also consider 'self-financing principle', i.e., the case in which the revenue from selling the permits to the road users is used for financing the capacity expansion of the network. Sakai et al. (2017) design Pareto-improving permit-based scheme for a V-shaped two-to-one merge bottleneck. They formulate the morning commute model in the network and describe the arrival time choice equilibrium in the network with merging bottleneck and show the conditions under which show that the first-best pricing scheme can achieve a

Pareto-improvement. They propose different implementations of bottleneck permits for Pareto-improving, and derive their equilibrium solutions for each implementation. Specifically, they demonstrate that when the permit revenues are used for expanding the bottleneck capacity, the Pareto-improvement state is achieved and social cost is decreased.

2.4.7 Parking Space Regulation

The roadway-use scheme can be used for managing parking space in areas with inadequate parking space. For instance, different permit-based schemes are studied in Zhang et al. (2011) for a many-to-one networks where each origin is connected to the destination by a highway with a bottleneck and a parallel transit line. Liu, Yang, Yin and Zhang (2014) develop a permit-based scheme for parking reservations implementable when commuters are either homogeneous or heterogeneous in their VOT. To resolve undesirable benefit distribution among commuters, they propose an equal cost-reduction distribution of parking permits where commuters with higher VOT receive fewer permits. Considering the parking space constraint at destination, Xiao, Liu and Huang (2019) propose two permit-based schemes for managing parking issue under three alternative modes, i.e., transit, driving alone and carpool. They investigate equilibrium state and system-optimal distributions of parking permits in a many-to-one multi-modal network setting. It is found that the prices of parking permits decrease with the parking supply, and carpoolers pay less for parking than a solo driver. Furthermore, for solo-driving and carpool vehicles, the undifferentiated permit scheme with a uniform price is more efficient than the differentiated permit scheme. However, when the parking supply is relatively low the undifferentiated permit scheme significantly changes the permit-holding order of solo drivers and carpoolers (Xiao, Liu and Huang 2019).

2.4.8 Transport Networks with CAVs

A rapid transition is taking place in the transportation sector with emerging concepts of mobility marketplace—a shared economy with a collection of marketable commodities. The basic premise is to treat the roadway transportation capacity as a collection of commodities or services that can be bought from the transportation market. This concept is increasingly becoming a reality with the technological developments in automotive industry. CAVs are one example of these technological advancements. More explicitly, communication, connectivity, and automated technologies installed in CAVs are main vantage sources

that in many ways will fuel this seismic shift in urban mobility system (Wang, Peeta and He 2019). First, vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), and vehicle-to-anything (V2X) communication technologies will enable a seamless connectivity, and thus enabling integrated monitoring and actively influencing the mobility marketplace. Second, CAVs will enter the marketplace in different forms including individually-owned automated cars, robotaxis, and large autonomous taxis. Third, a wide range of microtransit and mobility services, such as ride-sharing and ride-hailing, will emerge in the mobility marketplace. Forth, an advanced traveler information system will create a technology-informed mobility marketplace and thus enables CAVs to obtain more accurate information about travel paths compared to the conventional vehicles. Fifth, a centralized traffic management could leverage the connectivity advantage of CAVs, enabling new opportunities of implementing some advanced traffic management schemes in the mobility marketplace. A rich host of mobility options is foreseeable as a result, ranging from smart jitneys and robotaxis to autonomous ride-sharing shuttles (Cervero 2017, Babones 2018). These transformations will inevitably shape the mobility marketplace—a collection of commodities manageable by incorporating some market-driven scheme to control traffic and harmonize travel demand with respect to the roadway space (Yang and Wang 2011, Akamatsu and Wada 2017, Chen et al. 2020, Luo et al. 2019, Wang, Peeta and He 2019). In general, CAVs, if operated under a centralized traffic management system, might be managed as a fleet of a marketplace similar to those of airliners, and thus operate (or behave) according to a desired system objective.

Traffic control and travel demand management approaches for conventional traffic have been investigated for decades. In the decades to come, however, transportation network and traffic flow will experience a seismic shift due to the imminent advent of CAVs, which are prospected to enter the public streets in the next decade (Bertoncello and Wee 2015). A long transition period, however, is expected and needs to be planned for the decades to come. We will have a mixed-fleet transport environment composed of both human-driven and connected automated vehicles. In other words, we will continue having traffic flow of human-driven vehicles (non-CAV users) while observing increasing traffic flow from CAV users. A mixed-fleet transportation environment will present complex challenges for employing any regulatory tool to manage travel demand and control congestion. This is mostly due to the characteristics of non-CAV and CAV users, and their dynamical interactions. It will be even more complicated when the non-CAVs and CAVs share the same transport network under a credit-based scheme in their daily commutes. For example, the non-CAV users must decide between being charged for accessing less congested roadways or

spending more time in congested roadways for making routing decisions under uncertainty. On the other hand, CAV users will have the advantage of making more informed routing decisions and at the same time could be charged less or even subsidized under coordinated routing decisions. Moreover, with a proper incentive program a transport authority can speed up the adoption of CAVs. This will be in line with previous findings that a system-optimal traffic pattern can be achieved only with joint implementation of an advanced traveler information system and a proper traffic control scheme (Yang 1998).

Recently there has been increasing interest to develop new traffic control models, considering a mixed transport setting with CAVs, with the goal of proper incentive program design, efficient traffic control approaches and algorithms for this new setting (Zhang and Nie 2018, Li, Liu and Nie 2018, Chen et al. 2019). In this regard, Zhang and Nie (2018) explore the trade-off between efficiency in terms of the travel time savings and control intensity in terms of the number of automated vehicle users that let the central transport authority decide for their route. It is assumed that the controlled vehicles are guided in order to minimise the total travel time while the uncontrolled vehicles attempt to minimize their own travel cost. In the same spirit, Li, Liu and Nie (2018) look into stability and efficiency issues and how to bring under the control a mixed traffic system including both human-driven and autonomous vehicles to an equilibrium. It is still an open question whether or not the existing travel demand management models and schemes somehow can be used or extended to bring a mixed-fleet traffic setting under the control.

A few studies have been conducted to examine how to regulate traffic flow or distribute travel demand across road networks with CAVs. Advanced traffic management schemes have been recognized as one of the sources of the efficiency gain from CAVs to improve the throughput of transport facilities (Luo et al. 2019). In other words, with the integration of an ATIS and a traffic management scheme the possibility of actively motivating users to behave more in a desired manner is increasingly becoming a reality. This combination is relevant practically, providing more flexibility to manage travel demand and shape general public's habitual travel behavior towards a system-level goal such as maximizing social welfare, promoting desired traffic condition, and sustaining emission control programs. This is also in line with the basic premise that with joint implementation of emerging transport technologies such as ATIS and CAVs and proper traffic demand control schemes we can actuate a system-optimal traffic pattern. In general, if we can influence the travel decisions of roadway users, we can then manage and operate our road networks centrally and maximize the improvement potential offered by CAVs.

A traffic stream with a mixture of cooperative and non-cooperative users, however, will present complex challenges to bring the ever-increasing travel demand under control due to their operational characteristics, route overlapping and dynamical interactions of its users when sharing the same transport infrastructure. Recently the issue of travel demand management of a mixed transport environment has been fueled by the fact that in the coming decades our roadway vehicle fleet will most likely be made up of a mixture of CAVs and non-CAVs. One of the main motivations is that CAVs could operate in a cooperative way to a meet up system objective, in contrast to non-cooperative behavior who adopt a self-serving routing principle (Chen et al. 2020). In the spirit of modeling mixed traffic equilibrium behavior, a large chunk of studies have been presented in the literature and various forms of mixed traffic equilibrium have been put forward since the pioneering work of (Haurie and Marcotte 1985).

One stream of research looks into properties and equilibrium conditions of a mix traffic flow setting. In this regard, Harker (1988) looks into different equilibrium conditions on networks with multiple OD pairs, where each OD pair can obey either cooperative and non-cooperative behaviors. In a similar spirit, Van Vuren et al. (1989) and Van Vuren and Watling (1991) model a mixed traffic equilibrium problem and assume that cooperative users are provided with a route guidance system operated by a central transport authority. They investigate the conditions for uniqueness and stability of the stochastic user equilibrium. Similarly, Yang (1998) studies multiple equilibrium behaviors under the adoption of an ATIS with the objective of reducing uncertainty with recurrent network congestion. He proposes a convex programming model and establishes the conditions for the existence, uniqueness and stability of the network performance and demand equilibrium for any given level of the market penetration rate of ATIS. In the advent of ATIS, Maher and Hughes (1995) also study the potential of such a system to drive a multi-user class stochastic user equilibrium assignment towards the system-optimal conditions. Similarly, Lo and Szeto (2002) study the concept of mixed traffic equilibrium in an elastic manner and develop a methodology to study the trade-off among conflicting objectives of service providers, users, and the traffic management agency from a route planning and guidance system. Yin and Yang (2003) consider a specific ATIS whose objective is to reduce drivers' travel time uncertainty with recurrent network congestion where the users might not always comply with the advice provided by ATIS. Yang et al. (2007) examine the existence and uniqueness of solutions in a mixed behavior network equilibrium model involving routing behaviors of user equilibrium, system-optimum and Cournot-Nash travelers where each traveler makes routing decision given the routing decisions of other travelers. Wang, Peeta and He (2019)

consider that CAVs can follow the leading vehicle with lower headways than conventional vehicles. They use results by Levin and Boyles (2016) who provide mixed-traffic volume-delay functions.

The next stream of related studies is about designing approaches and schemes to manage mixed-vehicle transport environment. In this spirit, Li, Liu and Nie (2018) look into stability and efficiency issues and how to bring under the control a mixed traffic system including both human-driven and autonomous vehicles to an equilibrium. In the same way, Zhang and Nie (2018) explore the trade-off between efficiency in terms of the travel time savings and control intensity in terms of the number of automated vehicle users that let the central transport authority decide for their route. It is assumed that the controlled vehicles are guided in order to minimize the total travel time while the uncontrolled vehicles attempt to minimize their own travel cost. Chen et al. (2020) study traffic streams comprise a mix of CAVs and selfish vehicles and develop a path-control scheme based on a linear program to achieve the SO state of the network by controlling a portion of CAVs. However, their assumptions on a deterministic traffic setting are restrictive due to the factors involved with operating AVs including sharing common network infrastructures (routes and links) with non-CAVs and AVs human-driver mode, vehicles reaction and response times. Any of these factors can contribute to operational uncertainties on AVs' driving behaviors such as speed profiles in regulating the traffic behavior of non-CAVs and optimally distributing traffic demand at large. Our review shows that traffic control models and theories for mixed traffic conditions have been rarely addressed in literature in part due to their inherent mathematical and theoretical complexities along with unforeseen early emergence of CAVs. As such, there has been a resurgent call to develop new traffic models, considering the CAV environment and possible effects and interactions within a mixed-vehicle traffic condition, with the goal of preparing incentive program design, efficient traffic control approaches and algorithms for this new setting (Zhang and Nie 2018, Li, Liu and Nie 2018, Chen et al. 2019).

2.5 Solution Methods and Computational Experiments

The formulations of roadway-use problem discussed previously are computationally intractable; as a result, most of the studies have focused on developing specific solution methods. For example, in Han and Cheng (2015a), the Lagrangian dual formulation of the model is used as an efficient solution algorithm. Han and Cheng (2015a) find that the

Lagrangian dual formulation is a continuously differentiable concave maximization problem sharing the same optimal solution with the original model, and that its gradient can be obtained by invoking the successive average method. According to these properties, a two-step convergent solution algorithm is developed such that the outer iteration is implemented by using gradient projection method with a predetermined step size sequence while the inner iteration is implemented by using the successive average method. Given the non-convex property of the bi-level model, Li, Ukkusuri and Fan (2018) propose the pattern search algorithm embedded with the projection method. Wang, Gao, Xu and Sun (2014b) develop a relaxation algorithm for a continuous network design problem with a credit scheme and equity constraints.

In Wada and Akamatsu (2013) the solution is based on an auction mechanism by reformulating the optimization problem into a master problem and a sub-problem and then applying Benders decomposition principle. For practical applications, heuristics and meta-heuristics are expected to be likely implemented as real-life problems become considerably more complex and larger in scale. The number of solution methods introduced in the literature has grown over the past decades and the computational capacity of computers has increased exponentially, enabling to solve large instances of the real-life problems. Recently, in Han and Cheng (2017) a heuristic sensitivity analysis-based algorithm is developed to solve the optimal design problem, with analytical expressions of gradient information of the SUE link flow and ME credit price. The objectives and feasible sets in the bi-level problems of Wang, Liu and Huang (2018) are both complicated, which makes it hard to solve them in polynomial time. Wang, Liu and Huang (2018) apply a modified genetic algorithm to obtain the optimal supply for the OD-based travel permits, or the optimal number of auto users under the OD-based tolling scheme. The modified algorithm in each iteration deletes the infeasible chromosomes and produces new solutions again if the crossover and mutation operations generate infeasible ones, until it generates feasible solutions. To the best of the authors' knowledge, Wada and Akamatsu (2013) is the only study that proposes an evolutionary approach for the allocation of network permits on general networks with multiple OD pairs. They propose to obviate path enumeration by introducing a column generation procedure and prove that the proposed mechanism is truthful and converges to a dynamic system optimal allocation pattern.

In terms of modelling and solution aspects an agent-based perspective allows the handling of large-scale problems and relaxes assumptions such as user homogeneity, which brings more flexibly to model complexity in the system like learning and interaction. Tian and Chiu (2015) turn to agent-based economic-transportation hybrid modelling and set up

an integrated agent-based evaluation platform to better predict travellers' route choice and trading behaviours. The platform is an iterative process that consists of policy making and travellers' behaviour modules such that each individual traveller carries his or her personal memory across. The goal of establishing this framework is to provide further intelligence to potential policy makers' decision-making process. The proposed platform is able to make use of both individual level microscopic behaviour data as well as aggregated traffic flow and market performance data.

Xu and Grant-Muller (2016a) propose a simulation framework to analyse the mode-choice of travelers in the traffic network before and after the implementation of a credit-based scheme. This framework is applied to the case of Beijing, China where it is demonstrated that the credit scheme is a promising policy to reduce the total vehicle-miles traveled in the traffic network. Similarly, Miralinaghi and Peeta (2016) use simulation to show that the minimum travel time objective does not always generate effective paths and that departure time switches across all user groups, especially for users with a high value of travel time, while a minimum emissions objective yields desirable behavioural adjustments. Besides, minimum travel time credit design does not always generate minimum carbon emissions in the network, especially in networks with complex OD pairs and paths.

For a multi-modal setting, Nie and Yin (2013) consider the mode and route choices using numerical examples under the credit-based scheme with real-life data (to the best of our knowledge, this is the first paper that does this). Most commonly used reasonably sized networks and benchmark models are X-shape, Ziliaskopolous network, Nyguen-Dupuis network, and Sioux Falls network, which are used in several studies to show the applicability of the proposed algorithms. For specifications of these networks see Suwansirikul et al. (1987), Wang, Xu, Grant-Muller and Gao (2018), and Li, Ukkusuri and Fan (2018). In most of the numerical experiments the Bureau for Public Roads (BPR) function is adopted to define the link travel times for the specified networks. The networks of analytical models are small and artificial though they provide useful and influential guidance. However, dealing with larger real world networks and scenarios requires efficient solution procedures. As a numerical example, Wada and Akamatsu (2013) use the Sioux Falls network, which has 24 nodes and 76 links and 528 OD pairs with specific physical conditions for each link (i.e., free-flow travel time, capacity), to demonstrate the convergence properties of the proposed mechanism in a realistic network.

2.6 Conclusions

As discussed in the previous sections, a large body of research has focused on developing and formulating different variants of roadway-use rights schemes—a futuristic instrument to mitigate traffic congestion and its negative externalities in urban areas worldwide. This in-depth review of the state-of-the-art methodological advances on this topic provides the basic constructs and theoretical and analytical aspects of the problem in designing the schemes under certain conditions. Summarizing the body of literature on different variants of the roadway-use schemes, we identified the following remarks and prospectives for further improvement and applicability of roadway-use right schemes:

Problem modeling: The formulations of a roadway-use scheme for the general setting is not a trivial task. Usually a separable and monotonically increasing link performance function is considered which assumes travel cost on the route is the sum of its link travel times. This simplifying assumption is used in most of the studies, except for Wang, Liu and Huang (2018) and Zang et al. (2018) where the link travel time is assumed to be flow-dependent and Han and Cheng (2017) where the objective function is to maximize the network reserve capacity. The modeling of the roadway use schemes mostly involves a mathematical programming approach where the problem is formulated as optimization or variational inequalities, incorporating equilibrium conditions for permits or credits. Under convexity property, the KKT conditions can be used to investigate the pertinent theoretical properties and conditions under which equilibrium for users and optimum market prices can be sustained. Heterogeneity of users and demand elasticity are challenging to capture and implement in mathematical models. Availability of perfect information and rational travelers for system optimal approach are the most basic used assumptions in mathematical formulation with equilibrium constraints which are not in accordance with real-world settings.

Endowment problem: For the mobility right endowment problem, we observe that most of the studies follow a free and uniform distribution approach for allocation of mobility permits or credits. A uniform distribution scheme is in fact not sufficiently fair (Wang et al. 2012). Though the free allocation scheme can increase its social acceptability, it may not satisfy the system planner's efficiency and return expectations. On the other hand, a dictatorial approach can lead to inequality, entry barriers of potential (prospective) entrants and the receivers' eligibility verification

issues (Fan and Jiang 2013). The majority of these schemes incentivize users by allowing them to trade their mobility rights or redeem their unused rights. Some studies, however, propose a reward and charge mechanism in which those who travel during off-peak hours or use less congested roadways are rewarded and those who travel during congested time periods or use more congested roadways are charged (Nie and Yin 2013, Liu and Huang 2014, Nie 2015, Xiao et al. 2015, Zhu et al. 2017, Xiao, Long, Li, Kou and Nie 2019). Roadway-use right endowment requires having a well-designed system for periodic distribution and collection of credits or permits by the transport authority. A remedy to this is recycling credits charged (i.e. positive credit rate) on some travelers into subsidies (i.e. negative credit rate) given to others (Xiao, Long, Li, Kou and Nie 2019).

Pricing problem: For the mobility right pricing problem, almost all of the studies assume that the credit or permit price is determined within a free market. Among credit-based schemes, a few studies assume that the credit price can only be determined or changed by the transport authorities (Gao et al. 2018, Li, Ukkusuri and Fan 2018). The pay-for-use method relies on the existence of full information about the travel demand and cost functions which are difficult to obtain in practice (Wang, Yang, Han and Liu 2014). When users' valuations are uncertain, computing the near-optimal prices requires an efficient protocol to elicit pertinent information and determine the market price of mobility permits (Lessan and Karabatı 2018). A solution for the incomplete information issues is using auction-based reservation scheme, which is discussed in some studies (Akamatsu and Wada 2017, Su and Park 2015). In a congested time slot when the mobility demand from potential users exceeds the available capacity of the time slot, employing a mechanism that finds near optimal market prices for the permits or credits become important as it can help the transport authority to achieve a higher efficiency in terms of return of investment on top of mitigating traffic congestion.

Charging problem: Regarding the charging problem, our review shows that most of the proposed models consider the link-specific rates (charges) with few exceptions for OD-based and VMT-based charging schemes. Wang, Liu and Huang (2018) is among the works that proposes an OD-based charging scheme where travelers between the same OD pairs, independent of their used paths and links, are charged the same. The OD-based charging scheme enables more flexibility than other policies such as the link-specific ones to manage traffic (Wang, Liu and Huang 2018). On

the other hand, time-dependent charging schemes seem too complicated to be well received by users as they cannot predict the amount of charges for each specific link (route) in advance (Xiao et al. 2015). In a few studies, the charging mechanism has taken different forms, such as distance-based charging (Xu et al. 2018, Xu and Grant-Muller 2016a,b, Gao and Hu 2015) and time-varying charging (Li, Ukkusuri and Fan 2018, Wang, Wada, Akamatsu and Nagae 2018, Tian et al. 2013, Xiao et al. 2013). Generally, the time-based or time-place-distance-based (as for P&R and CBD, see Gao et al. (2018)) charges would enable more effective control of the travel demand on the network (Yang and Wang 2011). Our review shows that the proposed schemes have some differences and similarities in the attributes used with the existing proposed schemes. Their dependence on a periodic credit distribution and collection mechanism to sustain the credit circulation in the system is another operational issue that could incur administrative costs due to distribution or circulation of the credits. To circumvent this issue, the credit circulation is suggested that recycles credits charged (i.e. positive credit rate) on some travelers into rewards (i.e. negative credit rate) given to other travelers (Xiao, Long, Li, Kou and Nie 2019). Overall, recent findings show that reward-based travel demand management instruments can be an effective tool to manage traffic and promote adoption of emerging transport services (Tsirimpa et al. 2019, Zhu et al. 2019, Hu et al. 2015).

Implementation: The implementation of these schemes put some important administrative burdens on transport authority. This requires the transport authority to obtain detailed information of network links and a proper calculation of link-specific charges. Also, it can raise public concerns about the equity and fairness of the charging rates from the general public. The implementation of the link-specific schemes could become very complicated when a multi-modal and multi OD-pair traffic network with heterogeneous users and transaction costs are considered. A solution to this is to implement a second-best permit scheme. This (second-best) approach is actually less efficient but more practical as it can erase the negative effects induced by the OD-based and link-specific tolling policy (Wang, Liu and Huang 2018). The implementation-related issues, such as considering flexibility and efficiency, simplicity and fairness concerns in the design of these schemes are hardly touched. The implementation must deal with the government management agency, travelers, transportation firms, insurance companies, financial institutions and other intermediaries. In a similar spirit, institutional structures, operational rules for booking, cancellation, and trading introduce many challenges towards the practicality of

34

roadway-use schemes. In the same way, monitoring the usage of facilities in terms of time, places, and rules for detecting and fining non-compliances are other hardly touched operational issues (Fan and Jiang 2013). For a given set of transportation network characteristics, a transport authority must endow the users of the system based on their time-place-varying desires while satisfying efficiency measures in terms of level of service and return of investment.

Sustainability: To achieve sustainable development targets in a transportation market, equitable and yet efficient schemes need to be developed (Zhang and Waller 2018). A more realistic, however more complicated, approach is considering multiple objectives at the same time. One can consider maximizing the number of charge-free links, minimal charges or rewards, permissive charges, or minimizing excess travel time and other transport externalities such as total number of traffic accident fatalities, total network emissions, or fuel consumption as a secondary objective along with the common system travel time or cost objective function. The trade-off between different objectives, such as allocative efficiency, incentive compatibility, and budget balance should be evaluated. A co-evolutionary approach can be a good choice to model the interactions between different objectives and policy makers, technologies, institutions, business strategies and users (Foxon 2011). An extension to the (second-best) existing methods can be developing cordon-based, area-based, or distance-based supplementary schemes. Another extension can be considering a controlled congestion state in which the total transportation cost is minimized while a certain level of service is satisfied.

Prospective: Our review shows that potentials of roadway-use right schemes under ATIS for managing travel demand of a mixed-fleet setting have not been tackled so far in the literature in part due to their inherent mathematical and theoretical complexities along with unforeseen early emergence of CAVs. An operational issue is the existence of a free trading market for trading credits or permits. It can be realizable, for example, with the inclusion of a credit-base scheme within an ATIS in parallel with emergence of CAVs. This, however, motivates further studies to look for designing traffic control approaches with possibility of credit- or permit-based scheme and ATIS integration for traffic management and employing incentive programs and policies that can also act as a platform for trading credits or permits.

In our path to designing a mobility scheme, it is important to have an integrated architecture that address some or all of the identified issues above and considers users' and

operators' expectations, and traffic control and network requirements, in order to operate the entire system efficiently. This helps us to incorporate different components required for implementing a coordinative mobility scheme, take into account the influence of the participating players and the underlying issues. To this end, in the Chapter (3), we present the architecture of an integrated framework that embeds our proposed schemes. We then elaborate our design with its key components while addressing the expected challenges. Our review of the relevant literature suggests that existing models for roadway-use right schemes have not yet incorporated innovative designs and charging schemes on the basis of the mechanism design paradigm. A careful mechanism design practice helps relax many practical challenges such as the assumptions of users' rationality and existence of perfect information. It can also obviate issues related to the misreporting of spatial-temporal trip costs by travelers to reduce their generalized travel costs (Ren et al. 2020). Therefore, there is a need for careful mechanism design practice integrated to these schemes in order to solicit truthful information. To complete this gap, we contribute to the literature by proposing a permit-based scheme for single-bottleneck roadways setting and a credit-based scheme for general transportations networks in Chapters (4) and (5), receptively, using a mechanism design approach. We also present the theoretical properties of the proposed schemes and provide the results of comprehensive numerical experimentations under different settings. The proposed schemes come with right transferability option that enables redistribution of marginal welfare among users.

In Chapter (4), we design and analyze a mobility permit pricing and allocation on single-bottleneck roadway. Our review of the relevant literature suggests that existing models for permit-based traffic management have not yet incorporated innovative designs, such as user-centric and hybrid schemes, using the mechanism design paradigm. In this regard, there is room for improvement of the permit-based mobility management schemes by addressing application-related issues and concerns of the mobility users, the mobility service providers and authorities. To address these issues, we integrate a mechanism design practice and present a Pareto-improving permit-based mobility scheme which is applicable to single-bottleneck roadways with multiple involved stakeholders. Furthermore, we design a MP-based scheme that is strategy-proof and efficient in finding effective market prices suitable for travelers.

In Chapter (5), we design a credit-based mobility management scheme that address the issues of mixed-fleet traffic flow for transportation networks with advanced connectivity technologies. Particularly, we propose a traffic management scheme applicable to the settings where the traffic stream is a mixture of cooperative and non-cooperative users.

We present a set of revenue-neutral credit-based mobility schemes on general transportation networks under stochastic environment and interaction effects attributable to this transport setting. Unlike the traditional credit-based schemes, our schemes use a combination of distinctive charges and rewards (subsidies) that are usable to lead the network traffic to a lower system cost, or promote the adoption of CAVs, avoiding periodic collecting and distributing mobility credits in the end. Specifically, we demonstrate that a (Pareto-improving) revenue-neutral charging scheme is still attainable on a transportation network composed of CAVs and non-CAVs. We also study some specific circumstances that the proposed model can be applied with a few adjustments. Apart from the most of the proposed schemes, we consider a mixed CAV and non-CAV stochastic network flow setting. There are also differences in the protocols used in the proposed scheme, including a charge and reward strategy and a revenue-neutral policy, to address issues related to public acceptance and avoid difficulties in most current studies that are mainly based on assuming a charging scheme with a periodic credit distribution process.

Table 2.1 Summary of State-of-the-art Models

Citation	Class Type	Users Type	Right Endowment Type	Charging Type	Model Type
Yang and Wang (2011)	Credit-based	Homogeneous	All eligibles	Link-specific	Static
Zhang et al. (2011)	Permit-based	Homogeneous	All eligibles	OD-specific	Static
Wang and Yang (2012)	Credit-based	Homogeneous	All eligibles	Link-specific	Static
Wang et al. (2012)	Credit-based	Heterogeneous	All eligibles	Link-specific	Static
Wu et al. (2012)	Credit-based	Heterogeneous	All eligibles	Link-specific	Static
Nie (2012)	Credit-based	Homogeneous	Auction/Negotiation	Link-specific	Static
He et al. (2013)	Credit-based	Homogeneous	All eligibles	Link-specific	Static
Nie and Yin (2013)	Credit-based	Homogeneous	Charge/Reward	OD-specific	Static
Xiao et al. (2013)	Credit-based	Heterogeneous	Various	Time-varying	Dynamic
Tian et al. (2013)	Credit-based	Heterogeneous	All eligibles	Time-varying	Dynamic
Wada and Akamatsu (2013)	Permit-based	Heterogeneous	All eligibles	Link-specific	Static
Ye and Yang (2013)	Credit-based	Homogeneous	All eligibles	Link-specific	Dynamic
Shirmohammadi et al. (2013)	Credit-based	Homogeneous	All eligibles	Link-specific	Static
Bao et al. (2014)	Credit-based	Homogeneous	All eligibles	Link-specific	Static
Wang, Gao, Xu and Sun (2014b)	Credit-based	Homogeneous	All eligibles	Link-specific	Static
Wang, Gao, Xu and Sun (2014a)	Credit-based	Homogeneous	All eligibles	Link-specific	Static
Zhu et al. (2014)	Credit-based	Heterogeneous	All eligibles	Link-specific	Static
Gao and Sun (2014)	Credit-based	Homogeneous	All eligibles	Link-specific	Static
Li and Gao (2014)	Credit-based	Homogeneous	All eligibles	Link-specific	Static
Wang, Yang, Han and Liu (2014)	Credit-based	Homogeneous	All eligibles	Link-specific	static
Liu and Huang (2014)	Credit-based	Homogeneous	Charge/Reward	Link-specific	Static
Liu et al. (2015)	Permit-based	Heterogeneous	Auction-based	OD-specific	Static
Gao and Hu (2015)	Credit-based	Homogeneous	All eligibles	Per Kilometer	Static
Nie (2015)	Credit-based	Homogeneous	Charge/Reward	OD-specific	Static
Han and Cheng (2015b)	Credit-based	Homogeneous	All eligibles	Link-specific	Static
Han and Cheng (2015a)	Credit-based	Homogeneous	All eligibles	Link-specific	Stochastic

Table 2.1 (Continued) Summary of State-of-the-art Models

Citation	Class Type	Users Type	Right Endowment Type	Charging Type	Model Type
Tian and Chiu (2015)	Credit-based	Heterogeneous	Various	Link/OD/VMT*-based	Dynamic
Xiao et al. (2015)	Credit-based	Heterogeneous	Charge/Reward	OD-specific	Static
Xu and Grant-Muller (2016b)	Credit-based	Homogeneous	All eligibles	Per Kilometer	Static
Xu and Grant-Muller (2016a)	Credit-based	Heterogeneous	All eligibles	Per Kilometer	Static
Gao et al. (2016)	Credit-based	Homogeneous	All eligibles	Mode-specific	Static
Bao et al. (2016)	Credit-based	Homogeneous	All eligibles	Link-specific	Static
Shirmohammadi and Yin (2016)	Credit-based	Heterogeneous	All eligibles	OD-specific	Static
Xu et al. (2018)	Credit-based	Homogeneous	Per Household	Per Kilometer	Static
Han and Cheng (2017)	Credit-based	Homogeneous	All eligibles	Link-specific	Stochastic
Akamatsu and Wada (2017)	Permit-based	Heterogeneous	All eligibles	Link-specific	Static
Liu and Nie (2017)	Credit-based	Heterogeneous	All eligibles	Link-specific	Static
Zhu et al. (2017)	Credit-based	Heterogeneous	Charge/Reward	Link-specific	Static
Bao et al. (2017)	Credit-based	Heterogeneous	All eligibles	Link-specific	Static
Zang et al. (2018)	Credit-based	Homogeneous	All eligibles	OD-specific	Static
Wang, Gao and Xu (2019)	Credit-based	Homogeneous	All eligibles	Link-specific	Static
Gao et al. (2018)	Credit-based	Homogeneous	All eligibles	Link-specific, P&R*, CBD*	Static
Wang, Wada, Akamatsu and Nagae (2018)	Permit-based	Heterogeneous	All eligibles	Time-varying	Static
Li, Ukkusuri and Fan (2018)	Credit-based	Heterogeneous	Various	Time-varying	Dynamic
Wang, Liu and Huang (2018)	Permit-based	Heterogeneous	All eligibles	OD-specific	Static
Lahlou and Wynter (2017)	Credit-based	Heterogeneous	All eligibles	Mode-specific	Static
Miralinaghi and Peeta (2016)	Credit-based	Homogeneous	All eligibles	Link-specific	Multi-period
Wang, Xu, Grant-Muller and Gao (2018)	Credit-based	Homogeneous	All eligibles	Link-specific	Static
Xiao, Long, Li, Kou and Nie (2019)	Credit-based	Homogeneous	Charge/Reward	Link-specific	Static
Xiao, Liu and Huang (2019)	Permit-based	Homogeneous	Various	P&R*	Static
Sakai et al. (2017)	Permit-based	Homogeneous	All eligibles	Link-specific	Dynamic
Nie (2017b)	Permit-based	Homogeneous	All eligibles	Link-specific	Static

* VMT: Vehicle Mile Travel, P&R: Park and Ride, CBD: Central Business District.

39

Chapter 3

User-centric Paradigm on Designing Mobility Management Schemes [†]

As discussed in the previous chapters, it is important to deploy an effective mechanism to coordinate the travel demand of commuters throughout the mobility network, control traffic, and operate the entire system efficiently. This chapter presents the architecture of an integrated mobility management system in which our proposed permit-based mobility management is embedded. This helps us understand different components required for implementing a coordinative permit-based traffic management system, the influence of the participating players, and the underlying issues. We then elaborate our design with its key components and present our methodology for addressing the expected challenges.

3.1 Integrated MP-Based Traffic Management

The conceptual framework of the proposed MP-based traffic management system is illustrated in Figure (3.1). Different from Gärling et al. (2002) and Dogterom (2017), we take a holistic approach by considering the needs and functions of all actors (e.g., regulator, planners, providers and users) of a permit-based mobility management system. To implement such a coordinated MP-based mobility management mechanism, a futuristic mobility system is envisioned with the following assumed settings:

- Mobility becomes a market-based economy; users need to pay to use mobility services of any mode with specific mobility permits at competitive market prices,

- Mobility becomes a new type of service mostly offered by advanced demand responsive mobility providers, such as ride-hailing or ride-sourcing services.

- At transportation market, the competition is on offering users with personalized and competitive mobility options,

- Mobility network infrastructures become more and more integrated, and connections between different modes become seamless,

- Private vehicle usage decreases; people are expected to change their behavior from making conventional trip decisions - to ride here or there - to utilizing permit-based and point to point mobility-on-demand services,

- Vehicles are equipped with advanced tracking and communication technologies,

- Mobile Apps or other systems help mobility users in facilitating permit-based mobility decisions, providing them with real-time, place-dependent travel information,

- Mobility permit charges are tied to the type of service, time, ownership, and other relevant factors which are differentiated with specific time-place-mode dependent plans.

As depicted in Figure (3.1), the MP-based system works at the interconnection of mobility users, mobility permit and service providers, under the governance of a system regulator, for meeting the travel needs of users. One of the factors that influences the demand for mobility is the purpose of users for satisfying certain needs or obligations. Usually, users are faced with different service options; however, their choices and decisions on using any of the available options are influenced by their preferences, the status and attributes of the options, and the spatial, temporal and interpersonal factors. In the proposed mechanism, each user's preferences are translated into personalized off-line or on-demand service offers based on the information obtained from the user. Users are charged through permits or access rights whose prices are determined by a board of trustees, or ideally, in a free mobility market. In the market, travelers are allowed to trade their permits with each other freely under the regulator's supervision. To prevent a high permit price when the demand for mobility surges the system operator can hold a reserved capacity in order to respond to unexpected demand surges or supply shortcomings.
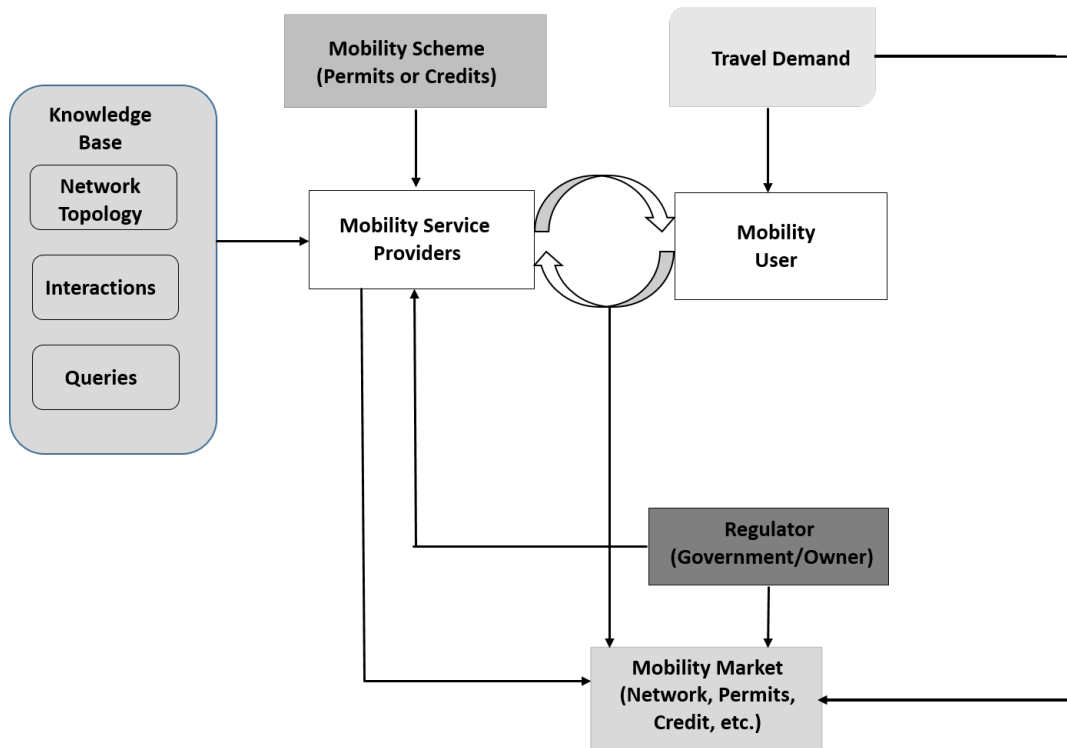
Figure 3.1 High-level Framework of the MP-based System

Figure (3.2) illustrates the main procedure of mobility permit endowment operation. In the first step, a front end user sends a permit request to the server of the selected service provider. A permit request includes an OD pair, a desired departure/arrival time, and a service type (on-demand or planned), and a mode of service (private car, taxi, carpool, and minivan). In the second step, the system offers a menu which consists of different service options, depending on the user's preferences. In the third step, the user chooses a subset of the options and confirms it with the service provider. The user may choose only one option or reject all the offers. In the last step, the server sends a notification/confirmation to the user and specifies the option he/she is assigned with.

For simplicity, but without loss of generality, we assume that the users are willing to follow the system instructions and proposals, e.g., option, route and time recommendations, when they are assigned with permits. While it is commonly recognized that users are self-concerned, we assume that a monitoring and penalization mechanism is implementable with emerging abilities to control trajectories of moving objects. Also, monitoring, scheduling, and routing operations require the system operator to enforce all users being equipped with a tracking device. Current communication and location-aware technologies allow tracking of the trajectories of moving objects with a high precision which can then be used to
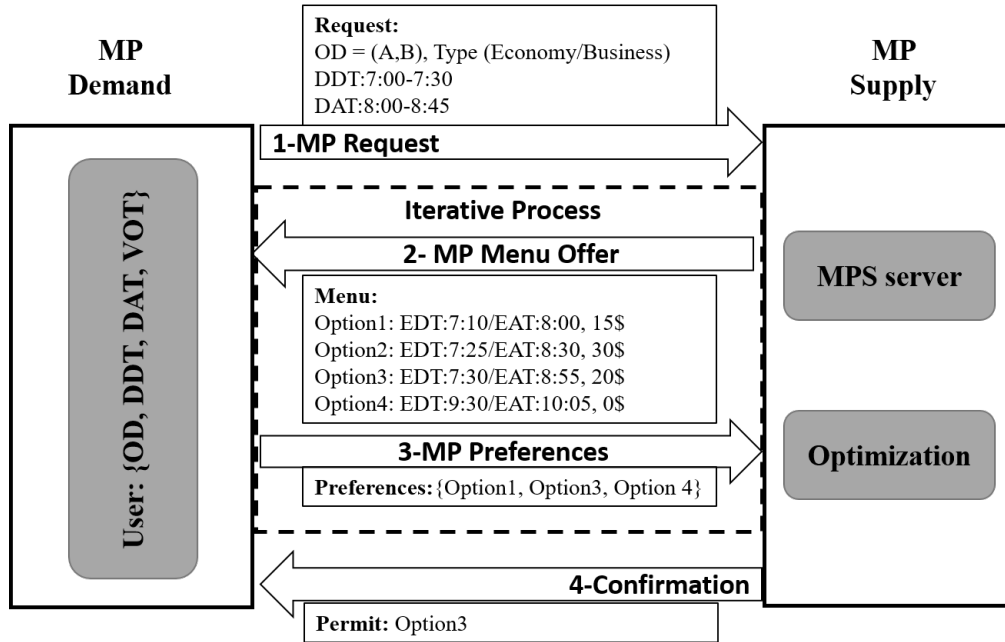
Figure 3.2 A Schematic of MP-based Reservation Process

provide mobility service recommendations and optimize traffic patterns based on real-time trajectory of users. However, this raises privacy concerns, for example, about the users' private information about their location. Therefore, the system regulator needs to put strict rules on mobility service providers to ensure that users' information will be held secure and only used to help operate and coordinate the system. The independence of the system regulator and competition of the service providers help encourage the transparency, fairness, and competitiveness of operations within the mobility market.

We have presented the procedure and process of the MP-based mobility scheme. The procedure for a credit-based scheme can be similar but with some operational differences. For instance, instead of permits, there are charges tied to each link of the transportation network about which users are informed and charged upon passing through them. The decision on whether all travelers or part of them have to participate is a managerial issue which is out of the scope of this research. We need to mention this decision also depends on the problem setting. For example, in a single-bottleneck network setting only travelers who want to pass through the bottleneck are involved in the permit or credit system. However, in a general transportation network setting with multiple bottlenecks the transport regulator may require all of the travelers to participate. In general, the users are faced with multiple options when traveling from their origin to destination, regarding routes and travel time and assumed to be flexible in choosing one of the several options between their

OD pairs. Specifically, individual users are expected to pick their routes based on their preferences which are often unknown to the system operators and could vary by location and time. Indeed, this is one of the difficulties for the mobility manager that the users' spatiotemporal utilities are uncertain. Under this situation, and to overcome the uncertainties associated with the users' private utility and their bounded rationality, it is necessary to integrate their expressed preferences over services. To this end, for each individual user's preferences and mobility service requirements, the proposed scheme provides them with flexibility to designate their preferences and priories from a menu that includes a list of affordable mobility options (services). To operate such a system, the main operational challenges are: how to set prices and allocate link-specific mobility permits over a transportation network; how to make users comply with such a system; how to meet the system planner (operator), service provider and regulator's expectations on the return of investment and operational efficiency. Next, we characterize the potential influencing factors for practicality of a MP-based mobility management system and discuss our solutions for the problems.

3.2 Challenges and Proposed Solutions

The novelty of the proposed permit-based scheme discussed in Section (3.1) also comes with several technical challenges, including personalization and customization, simplicity, computational efficiency, user-acceptance, and system stability. This section presents some ideas on how to address these challenges.

3.2.1 Personalization and Customization of Permits or Credits

The customization of permits or credits are one of the distinctions of this work from the other proposed schemes in the literature. For example, using a permit reservation medium, the mobility system manager collects users' mobility preferences such as the origin and destination pair, desired departure time (DDT), desired arrival time (DAT), desired speed, acceptable earliness and lateness, mode of service (economic, business, emergency, minimum time, or congestion free). Based on the availability of permits and the announced permit prices, each individual user can sort out a short list of the preferred service options based on their utility. The preferences are expressed as an ordering of the provided options to the service provider. Each service option corresponds to a bundle of link-specific permits

for the set of links required to use that service. We assume that each user has a time-dependent private valuation for each option of arriving to the destination with scheduled earliness, lateness, or at desired time period. Personalization of mobility permits has not been addressed in the previous schemes. For example, Wada and Akamatsu (2013) proposed a MP-based mechanism under which each commuter needs to purchase a number of permits depending on their preferred path. However, in practice it is cumbersome for users to figure out a bundle of permits out of available options for their desired mobility services. To solve this issue, in the proposed scheme the service provider offers a service menu based on each users' OD pairs and desired arrival/departure time. Listing different mobility options, a service option in the menu is a combination of: 1) expected arrival time at the destination, 2) total permit cost to pass the links (bottlenecks) of each route during the specified time. This is one of the distinctions of the proposed scheme to reduce the complexity of the MP-based mobility management system that is found to have a large impact on the public support. In the case of credit-based scheme, we design credit charges specific to users' types and also roadways (links) of the transportation network. To avoid a long discussion here, we will discuss the details of credit customization in Chapter (5).

We relax the assumptions on users' rationality and availability of full information and assume that users can have any sort of mental processes (to maximize their private individual utility or behave in a completely irrational way) in determining their own preference list. Indeed, by letting users choose and announce their own preferences out of the available mobility service options, we can overcome the conventional restrictive assumption on users' rationality and uncertainty about their utilities. User heterogeneity is usually taken into account by considering a set of discrete VOT functions or by defining continuously distributed VOT functions (Wang et al. 2012); however, such functions are difficult to be established. Different from the literature, in order to cast user heterogeneity effect we propose integrating users' choice behavior by letting them have their own preferences throughout the permit endowment process. Therefore, we assume that the system manager provides users with multiple permit options and allows them the flexibility to choose their preferences out of the provided options. The manager ideally seeks to solve the system optimal mobility permit allocation problem. Although the system may not guarantee assigning every user to their most preferred requested service, it assigns them to their highest preferred option using a permit allocation scheme.

3.2.2 Allocation of Permits or Credits

Center to a MP-based traffic management system is a component solving a resource allocation problem, in which the mobility service provider decides on the allocation of scarce resources (permits and usage rights) to users. Therefore, efficient allocation of mobility permits has the highest priority, which should be carried out in a sequential steps. This task, however, depends on several factors such as OD demands, network characteristics (bottlenecks and links dependencies), alternative options, background traffic, and considered time periods.

First, in a sequential manner, big time slots such as months and weeks are divided into several shorter time slots such as days and peak and off-peak hours. Each time slot is allocated with specific number of mobility permits given the predicted traffic pattern at network/roadway scale. Mobility demand dynamics over the entire transportation network is another key factor, when designing a MP-based traffic management scheme. In this regard, we consider short-term planning horizons during which the capacity of the network is fixed, and assume that the travel demand during the horizon is stationary and finite. To prevent the congestion issue over the entire network, the total allocated quotas to mobility service providers along with the withheld (reserved) quotas cannot exceed the network capacity. Moreover, to manage temporary or unexpected mobility demand surge or supply shortage, the system manager uses the reserved capacity.

Then, mobility users express their preferences and then participate in the proposed permit allocation mechanism. As discussed in the previous section, based on users' information and the availability of permit quotas permits are issued to endow the holders to fulfill their mobility need. Specifically, a mobility need is actually a trip from any origin to any destination of the permit-based mobility network that needs to use one or multiple bottlenecks of the network. Finally, based on the users' expressed preferences, the system operator assigns them with permits to use the roadway network such that the overall traffic pattern is optimized. We will show that the traffic pattern can be coordinated and optimized with well-defined permit allocation policies. We note that the permit trading market can also be designed and monitored by the system planner to allow the tradability of endowed permits and distribution of the welfare in the form of exchanging permit for permit, or money for permit and vice versa.

In the case of credit-based scheme, there are also differences in the protocols used in the proposed scheme. Particularly, a credit circulation policy through which is suggested that recycles credits charged (i.e. positive credit rate) on some travelers into rewards (i.e.
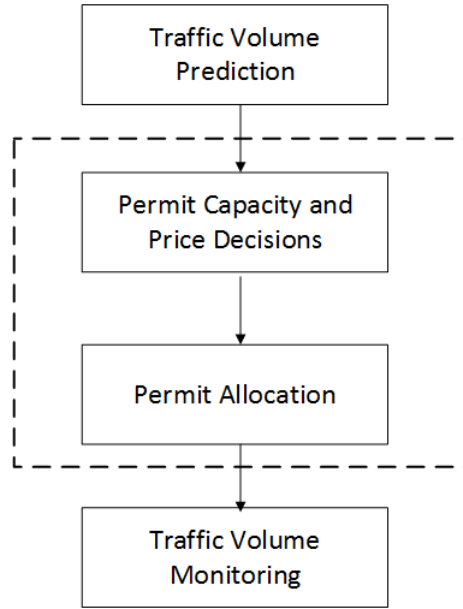
Figure 3.3 Components of a Permit-based Mobility Control System

negative credit rate) given to other travelers. In this way, we can avoid difficulties in most current studies that are mainly based on assuming a charging scheme with a periodic credit distribution process. To avoid a long discussion here, we will discuss the details of credit circulation policy in Chapter (5).

Figure (3.3) depicts components and steps to operate a mobility permit-based management scheme. As discussed, the allocation of the mobility permits are the main challenges, which influence the viability, effectiveness and acceptability of a permit-based traffic management system. In this research, we consider two different settings. First one is a single-bottleneck roadway setting, and the second one is a transportation network with multiple bottlenecks with several pairs of origins and destinations. In the either of these settings, a service provider must endow mobility permits at effective (congestion mitigating) market prices to users of the system based on their time-varying desires, travel needs and their expectations of fairness while achieving a satisfactory level of investment return. Specifically, in Chapter (4), using a mathematical programming approach, we show how these restrictions can be embedded in a constrained coordinated permit allocation model with efficiency and fairness objectives on top of observing different operational constraints.

3.2.3 Fairness and Efficiency

As discussed in the previous section, our focus in this study is on developing (near) optimal pricing and allocation of permits under a fairness requirement. Therefore, two operative goals must be taken into account in making mobility permit allocation decisions. The first one is allocative efficiency, a degree to which the entire system operates close to a system optimal state with respect to a collective measure from a central planner's point of view. The second one is equity or fairness, a degree to which the outcome of the scheme satisfies the desires and needs of each mobility user. The issue of balancing efficiency and fairness has been studied extensively in several traffic management literature; however, it has not been addressed in MP-based traffic management schemes.

Due to the subjective nature of fairness and different possible interpretations of equity, there is no unique "one-size-fits-all" solution for fairness (Karsu and Morton 2015), or a principle that is universally accepted as "the most fair"(Bertsimas et al. 2012). However, the most recognized measures and dimensions of fairness are, namely, "horizontal" and "vertical" equity (Levinson 2010, Viegas 2001) *. The "horizontal" concept is also called *equitability* which is concerned with the "equal treatment of equals". On the other hand, the "vertical" equity is a balance measure concerning with the "unlike treatment of unlikes", distinguishing the entities by their specific attributes such as their needs, claims or preferences[†]. Equitability is distinguishable from balance depending on whether an underlying anonymity assumption holds, i.e., it holds when all the permutations of users are treated indifferently (Karsu and Morton 2015).

Along with equity considerations throughout the permit allocation process, a system manager (regulator or policy-maker) is concerned with allocative efficiency (e.g., net throughput of the network as total system travel time) to meet the return of investment and revenue drivers. However, efficiency and equity are two conflicting measures in establishing mobility management systems. A system design resulting in the highest social surplus (including time gains, paid charge and adaptation costs plus revenues) is considered the most efficient yet the least equitable design (Kristoffersson et al. 2017). The efficiency concern has been a sole criterion in most of the conventional mobility manage-

---

*We do not consider other measures related to equity, such as envy-freeness.

[†]A horizontally equitable policy distributes treats equally all individuals. A vertically equitable policy, on the other hand, favors groups that are socially or economically disadvantaged, for instance low-income groups. The vertical equity is categorized into vertical with respect to income and social class, and vertical with respect to need and ability. These two types of vertical equity are typically evaluated through a welfare-based and access-based approaches, respectively (Litman 2002).

ment schemes. However, we consider both (efficiency and fairness) as key technical and practical issues in mobility management systems. One natural way, in the literature, to achieve an efficiency-equity trade-off is to use an aggregation function that reflects concerns for both equity and efficiency, which will be adopted in this research. By maximizing such aggregation (value) function, one can avoid favoring some entities too much while depriving some others. However, a drawback with an aggregation function is that the decision maker has to set *a priori* an aggregation scheme, which is not in accordance with the equity aim of the intended scheme and thus less likely to be perceived as a fair solution (Karsu and Morton 2015). Despite this challenge, we propose an alternative approach to tackle the restrictions imposed by the fairness-efficiency requirements. We integrate the vertical and horizontal equity measures into the mobility permit allocation process in order to deal with the trade-off between efficiency and fairness. This is another distinction of this work from the conventional approaches, to the best of our knowledge. In the case of our credit-based scheme, we suggest a charge and reward strategy and a revenue-neutral policy, to address issues related to fairness public acceptance. To avoid a long discussion here, we will discuss the details of credit-based scheme in Chapter (5).

3.2.4 Pricing of Permits or Credits

How to price mobility permits is another key issue in implementing a MP-based traffic management scheme. Specifically, in a congested time slot when the mobility demand from potential users exceeds the available capacity of the time slot, employing a mechanism that finds near optimal market prices for the permits can help the mobility service provider achieve a higher efficiency in terms of return of investment on top of mitigating traffic congestion. However, when users' valuations are uncertain, computing near-optimal prices requires efficient protocols to elicit pertinent information and determine the market price of mobility permits (Lessan and Karabatı 2018). Our literature review reveals that the lack of fairness and ignoring users' preferences are the common weaknesses of almost all the proposed mechanisms for pricing mobility permits. This is also due to the fact that the proposed schemes are built on system-optimality expectations of system planners. Some of the proposed endowment and pricing mechanisms follow a free permit distribution approach (i.e., uniform or grandfathering), and some other follow pay-for-permits allocation scheme. Though a free way of permit allocation scheme, for example, can increase its social acceptability, it does not satisfy the system planner's efficiency and return expectations. On the other hand, a grandfathering approach can lead to inequality, entry barriers of

potential (prospective) entrants and the receivers' eligibility verification issues (Fan and Jiang 2013). The pay-for-permits method relies on the existence of full information about the travel demand and cost functions which are difficult to generate in practice (Wang, Yang, Han and Liu 2014).

Some conventional allocation schemes such as first-come first-served (FCFS), queue-position, or priority-based are unable to differentiate commuters based on their VOT and or account for their schedule preferences which would lead to efficiency loss. In this study, we propose to apply a progressive auction approach with which the mobility service provider and the users go through an auction process to determine the allocation outcome and price of permits for each time slot. The mobility service provider plays the auctioneer role and starts with an initial price vector for the time slots, P_0, equal to the minimum sales prices (reservation prices). The users follow the auction by selecting their desired permits. At each iteration after performing a provisional permit pricing and allocation, given the reactions of mobility users to the posted prices, the mobility service provider decides to raise the permit price of one or more time slots to proceed the auction to the next iteration or stop it at the current iteration. The iterative process of provisional pricing and allocation continues until a stopping criteria is satisfied. When the decision is made to stop the auction, the allocation of permits are finalized based on the assignments generated on the last iteration of the auction and requirements on the efficiency and equity concerns. The termination criterion of the iterative auction mechanism relies on the objectives of the service provider, which in our case is efficiency maximization goal while observing fairness of the allocations. As users place monetary values on each of the offered options, thus, given a price vector, they choose those items that bring them maximum benefits. A price vector then yields equilibrium if every user can be assigned to one service option while no capacity limit is violated.

It is worthwhile to mention that our auctioning scheme uses an open-bid, multi-item and user-oriented pricing approach which is different from similar works such as Wada et al. (2010) and Wada and Akamatsu (2013) that implement sealed bids and single-item mechanisms. Different from Liu et al. (2015), we incorporate users' preferences in the our auction mechanism to find the effective market prices of time slots given the entire users' preferences while guaranteeing them to be allocated with one of their preferences. We will show the existence of an equilibrium, and the uniqueness of the pricing outcome under a properly designed auction mechanisms. We need to mention that in case of credit-based scheme, the price of permits can be determined through a market clearing state and that this price is scalable. Therefore, the issue of pricing credits is less challenging. To avoid a

long discussion here, we will discuss the details in Chapter (5).

3.2.5 Computational Efficiency

The MP-based pricing and allocation problems, once formulated mathematically, is expected to be computationally intractable which will be a key challenge to the permit-based traffic management scheme. There are various approaches for dealing with computationally intractable allocation problems, such as applying stochastic search methods such as randomization, using heuristics such as genetic algorithm, or relaxing some constraints of the problem. We focus on developing randomized algorithms for solving large size instances from real-world settings. Since the design of such algorithm depends on each specific problem and the assumptions made in this regard, we will provide detailed discussions about the approximation algorithms of single bottleneck cases in the next chapter.

3.2.6 Performance Analysis

As discussed before, we are dealing with two contrasting goals, namely, efficiency and fairness. It is hard to find a unique criteria to measure the performance of the proposed mobility permit-based scheme under these criteria. It is a problem of balancing trade-off between efficiency loss and fairness loss; the efficiency obtained from a fair solution in general would be less than the efficiency obtained from a system optimum solution, and the equity obtained from an efficient solution can be far less desirable based on the fairness objective. However, to evaluate the performance of the proposed MP-based schemes we introduce two different metrics to MP-based traffic management systems.

First, we introduce the Price of Fairness (PoF) from Nicosia et al. (2017) and Trichakis (2011), as a standard indicator to measure efficiency loss due to introducing fairness to the system, that is defined as

$$PoF(U;S) \;=\; \frac{SO(U^*) - UMPA(U;S)}{SO(U^*)}. \tag{3.1}$$

Where $SO(U^*)$ is the maximum efficiency from a system optimal allocation that maximizes the efficiency of the entire system (U), and $UMPA(U;S)$ is a solution (S) that observes fairness requirements through the permit allocation. PoF measures the relative

loss of utility due to a mobility permit allocation solution. This can give the decision maker a guideline about the cost of fairness (Nicosia et al. 2017).

Next, we then restate the price of efficiency (PoE) from Trichakis (2011), which is defined as the total welfare loss from a $UMPA(U; S)$ solution relative to the maximum obtainable welfare from an equity-oriented system ($ES(U^*)$). The PoE metric is defined as

$$PoE(U; S) \quad = \quad \frac{ES(U^*) - UMPA(U; S)}{ES(U^*)}. \tag{3.2}$$

These metrics which are specific to our permit-based schemes let the mobility planner observe the trade-off between efficiency and fairness in allocating mobility permits. Other than these metrics, we will perform rigorous sensitivity analyses to check the stability of the proposed schemes against the dynamics of mobility demand such as travelers' heterogeneity and loss aversion. We need to mention that in the case of our credit-based scheme, we perform our comparisons about traffic state against the so-called user optimal and system optimal conditions, similar to the existing literature.

Chapter 4

Mobility Permit-based Traffic Management Scheme for Single-bottleneck Roadways [†]

4.1 Summary

As discussed in details in the previous chapter, the pricing and allocation of mobility permits is an essential part of a MP-based traffic management system. As roads with bridges and tunnels are the most obvious examples of a roadway network with a single-bottleneck, we first propose a mobility permit allocation mechanism using a mixed-integer programming model for single-bottleneck roadways. We then introduce the steps of a price formation method based on a progressive auction mechanism. Next, we discuss the properties of the proposed scheme. Finally, we demonstrate the performance of the proposed MP-based scheme considering various scenarios and metrics under different settings.

Figure 4.1 Single-bottleneck Roadway Model

4.2 Allocation of Mobility Permits: Problem Formulation

4.2.1 Preliminaries

We build an optimization model for mobility permit allocation on single-bottleneck roadways based on the departure time choice problem (Liu et al. 2015), with operational requirements. After describing the discrete time traffic flow and congestion dynamics on a single-bottleneck roadway, we then formulate the user-centric mobility permit assignment model that provides an optimal traffic flow pattern with respect to users' expressed preferences. The single-bottleneck roadway model is a modification of the basic morning rush hour problem where a continuum N of heterogeneous commuters want to travel from an origin (O) to a destination (D) using a roadway with a single bottleneck of a fixed capacity (M), as shown in Figure (4.1). We do not make any restrictive assumptions on the users' travel information or utilities; however, we suppose that all prices and users' valuations are integers, and the number of users sharing the roadway space is large enough such that each user's behavior has infinitesimal effect on the other users. We assume that there are a considerable number of users (individual travelers or freight transporters) who have preferences with respect to the available services and they have enough flexibility over the start or completion time of their mobility desires. In other words, they are willing to accept scheduled earliness or delay costs if they cannot find an ideal service time. Users are assumed to be asymmetric regarding their spatiotemporal mobility desires and information and are self-concerned, i.e., they are concerned about their own travel and indifferent about the others.

Without loss of generality, we assume that the stationary travel demand during a time
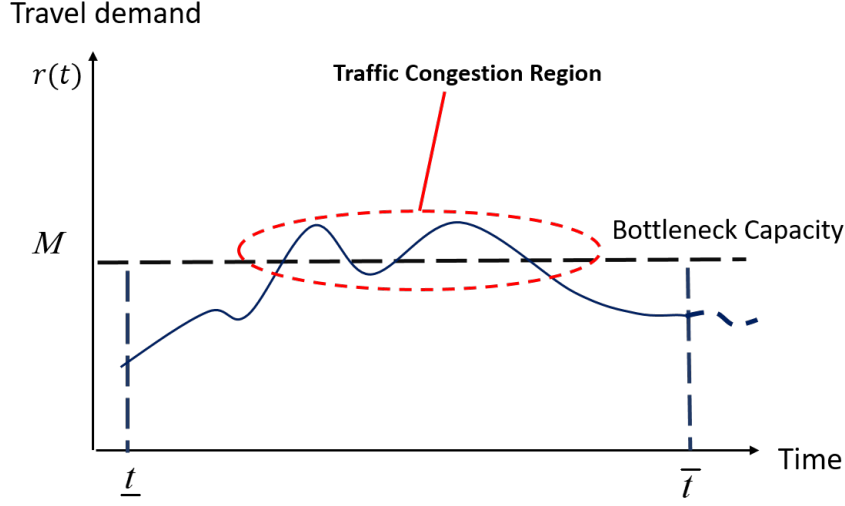
Figure 4.2 Traffic Dynamics in a Single-bottleneck Roadway

interval $[\underline{t}, \bar{t}]$ is finite and bounded above by $N \in Z_{>0}$, and the capacity of the bottleneck is fixed during a short-term horizon. Then, the following condition holds: $\int_{\underline{t}}^{\bar{t}} r(t)dt \leq N$, where $r(t)$ is the departure rate. Figure (4.2) depicts the mechanism of travel demand with respect to the roadway space (supply). Clearly, when travel demand reaches the supply (bottleneck) capacity, traffic congestion happens. To monitor and control mobility demand and thus traffic congestion, the regulator divides the interval $[\underline{t}, \bar{t}]$ into a set of chronologically increasing time points, i.e., $\{\underline{t} = t_1, t_2, ..., t_{K-1}, t_K = \bar{t}\}$, as shown in Figure (4.3), where $K \in Z_{>0}$ is the number of time slots. For each time interval $k \in K$ a continuum of potential users (N) of mass θ_k request passing the roadway bottleneck. However, to avoid congestion, the total number of users who are allocated with mobility permits to pass the bottleneck at each specified time slot should be less than the capacity limit (M) for that time interval.

The mobility service provider uses a reservation system to issue a limited number of travel permits with which permit holders are allowed to use the bottleneck within pre-specified time intervals. Specifically, the $k^{th} \in \{1, 2, ..., K\}$ interval dedicated to travel pass tp_k is denoted as: $tp_k = [t_{k-1}, t_k], 1 \leq k \leq K$. The commuters can make reservations for any of the travel passes. However, a reservation system that offers multiple choices (in the form of a menu) and allows commuters to express their preferences seems more preferable for users. To this end, the mobility service provider offers a menu $\boldsymbol{PM} = (\Delta, P)$ with $|\Delta| = |P| = L$ options where $\Delta = (tp_1, tp_2, ..., tp_L)$ is the set of available mobility options such that tp_l can be used to pass the bottleneck during l^{th} time interval $(tp_l \subset [\underline{t}, \bar{t}], l =$
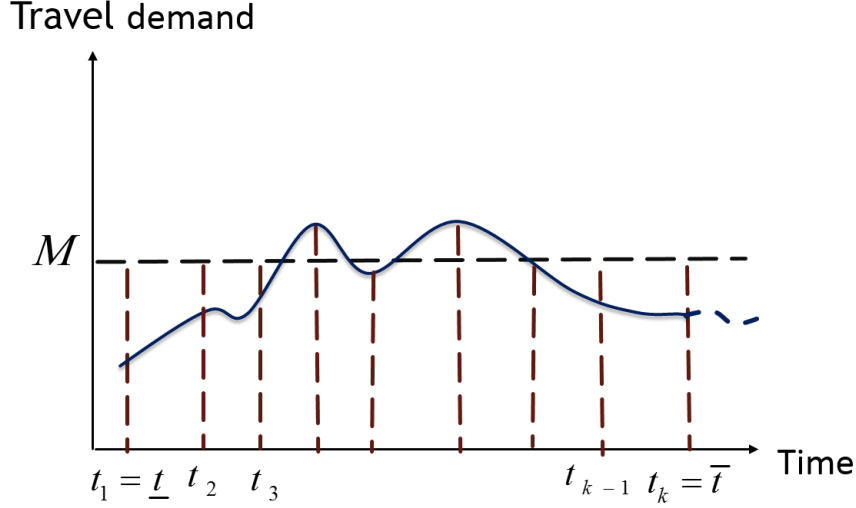
Figure 4.3 Pre-specified Time Intervals for Single-bottleneck Roadway Usage

$1, 2, ..., L)$ and $P = (p_1, ..., p_L)$ is the set of corresponding charges $(p_l > 0, l = 1, 2, ...L)$. The permits are differentiated with respect to time-intervals at different price levels depending on $t \in tp_l; l = 1, ..., L$.

Each user $i, i = 1, 2, ..., N$, is allowed to choose their desired options from the menu and specify their preferences. Let $U_i(tp_l)$ be the utility of buyer $i, i = 1, 2, ..., N$, when purchasing the lth option. We then let $\boldsymbol{q}_i = (q_{i,1}, ..., q_{i,L})$ be the preference list of user $i, i = 1, 2, ..., N$, where $\hat{q}_{i,l} = \arg\max_{k \in q_1, ..., q_L \setminus \{q_{i,1}, ..., q_{i,l-1}\}} U_i(tp_k), l = 1, 2, ..., L$. In other words, each commuter's decision problem is equivalent to searching and sorting through the offered permit options and then announcing their preference list. Although each user's travel information is private, they can be assumed to have monotonically decreasing order preferences, i.e., $U_i(q_{k-1}) \geq U_i(q_k)$ for $k = 2, ..., |L|$, based on a privately known strict ordering relation, "$\succeq_i$", which reflects user i's utility function of any form.

To avoid congestion, the number of accepted and finalized requests to pass through the bottleneck in the kth travel interval should be within the capacity of the bottleneck during that time interval, i.e., $M_k \leq \sum_i^N \delta_{lk}^i tp_l$, for $k = 1, ..., K$. The total travel capacity during $[\underline{t}, \overline{t}]$ is $\sum_{k=1}^L M_k > N$. The system manager can eliminate traffic congestion by limiting the number of allocated permits to be less than or equal to the bottleneck capacity of that time period. In other words, a congestion-free traffic pattern can be described with a vector of issued permits $\boldsymbol{y} = (y_1, ..., y_N)$ with $y_i = (y_{i1}, ..., y_{iL})$, such that $\sum_{i=1}^N \sum_{l=1}^L y_{i,l} \delta_{lk}^i \leq M_k, k = 1, 2, ..., L$, where δ_{lk}^i is the Kronecker delta (i.e., 1 if l^{th} preference of user i is the kth option in the price menu; and zero otherwise) and $y_{i,l} \in \boldsymbol{q}_{i,k}$ or $y_{i,l} = 0, i = 1, 2, ..., N$. The vector

$\hat{\boldsymbol{p}} = (\hat{p}_1, ..., \hat{p}_N)$ is the payment vector, where $\hat{p}_i, i = 1, 2, ..., N$, indicates the price that the user $i, i = 1, 2, ..., N$ incurs, if he/she is assigned with y_i. In particular, $p(t) : [\underline{t}, \bar{t}] \to \mathbb{R}$ is a step function which can be written as $p(t) = \sum_{l=1}^{l=L} p_l 1_{tp_l}$ where p_l are real numbers and 1_{tp_l} is the indicator function that takes the value of 1 if $t \in tp_l$ and zero otherwise.

In what follows, assuming a given price vector p, we first list the parameters and decision variables used for the mathematical programming of a user-centric permit allocation scheme with efficiency-equity notion, and then present the respective MIP model.

- Indices:
  i : is the index of commuters for $i = 1, 2, ..., N$,
  k : is the index of time slots for $k = 1, 2, ..., K$,
  l : is the preference index in a decreasing order for $l = 1, 2, ..., L \leq K$.

- Parameters:
  M_k: is the bottleneck capacity during k^{th} mobility pass interval, for simplicity $M_k = M$,
  $q_{i,l}$: is the lth preference (selected option) of commuter i,
  p_k: is the price of kth travel pass interval,
  δ^i_{lk}: is Kronecker's delta; takes value of one if the lth preference of commuter i is the kth option in the price menu; and zero otherwise
  BM: a large positive number (Big M).

- Decision Variables:
  $x_{i,l}$: a binary variable that takes the value of one if commuter i is assigned to the lth preference; zero otherwise.
  $z_{i,l}$: a binary variable that takes the value of one if the lth preference of commuter i is allocable at the corresponding travel interval; zero otherwise.

Using the above parameters and decision variables, the user-centric mobility permit allocation problem (UMPAP) is formulated through the objective function (4.1) and the constraint sets (4.2) to (4.8). This objective function finds the allocation with the highest possible efficiency (here in terms of monetary return) from users while assigning each user to their most-preferred travel path where it is available in the specified time interval*. This

---
*No individual user is guaranteed, *a priori* to be assigned with their first-choice by all circumstances.

model is a variant of the standard user equilibrium model with additional constraints on the permit allocation conditions.

$$\text{UMPAP:} \quad \max \quad \sum_{i=1}^{N}\sum_{l=1}^{L}\sum_{k=1}^{K} x_{i,l}\,p_k\,\delta_{lk}^{i} \tag{4.1}$$

s.t:

$$\sum_{i=1}^{N}\sum_{l=1}^{L} x_{i,l}\,q_{i,l}\,\delta_{lk}^{i} \quad \leq \quad M_k, \quad k=1,2,...,K, \tag{4.2}$$

$$\delta_{lk}^{i}=1 \quad \Rightarrow \quad \begin{cases} q_{i,l}+BM(z_{i,l}) \geq M_k, & i=1,2,...,N; \ \ l,k=1,2,...,K, \\ q_{i,l}+BM(z_{i,l}-1) \leq M_k, & i=1,2,...,N; \ \ l,k=1,2,...,K, \end{cases} \tag{4.3}$$

$$\sum_{l=1}^{L} x_{i,l} \quad = \quad 1, \quad i=1,2,...,N, \tag{4.4}$$

$$x_{i,l} \quad \leq \quad 1-\sum_{j=1}^{l-1} x_{i,j}, \quad i=1,2,...,N; \ \ l,k=1,2,...,K, \tag{4.5}$$

$$z_{i,l} \quad \leq \quad 1-\sum_{j=1}^{l-1} x_{i,j}, \quad i=1,2,...,N; \ \ l,k=1,2,...,K, \tag{4.6}$$

$$x_{i,l} \quad \leq \quad z_{i,l}, \quad i=1,2,...,N; \ \ l,k=1,2,...,K, \tag{4.7}$$

$$(1-z_{i,l})+x_{i,l} \quad \geq \quad x_{i,r}-(1-z_{i,r}), \quad i=1,2,...,N;$$
$$l=1,2,...,L; r=l+1,...,L, \tag{4.8}$$

$$x_{i,l}, z_{i,l} \quad \in \quad \{0,1\}, \quad i=1,2,...,N; \ \ l=1,2,...,L. \tag{4.9}$$

Constraint set (4.2) observes the capacity limit of the bottleneck for each travel pass interval k defined by the system manager. The constraint set (4.3) forces the variables $z_{i,l}$ to take the value of one if the lth preference (of type k) of commuter i is less than the available travel pass capacity of type k; and zero otherwise. We note that the constraint sets (4.4) to (4.8) are to check the eligibility of each user i to be allocated if none of their preferences with a higher priority has been allocated to their earlier. Through the constraint set (4.8), we implement equitability and balance requirements. In other words, it is guaranteed that if commuter i is allocated with a permit, the allocation is made for their top allocable choice. The user-centric allocation scheme within this model gives each user an equivalent access opportunity (equitability) at the level of anonymity by not differentiating the users in the selection step, after which it tries to balance the allocations by distinguishing users

given their (personal) preferences. Finally, to avoid partial allocations, we restrict the decision variables $x_{i,l}$, $z_{i,l}$ to be binary, using the constraint set (4.9).

The UMPAP is a coordinated constraint allocation problem that is an extension of the classical generalized assignment problem first introduced by (Ross and Soland 1975), which is here bounded by additional constraints to observe users' priorities. A similar idea of preference-oriented assignment was initially introduced by Lessan and Karabatı (2018) where it was used to find the pessimistic scenario of a randomized allocation problem.

Theorem 4.2.1. *User-centric mobility permit allocation problem is $\mathcal{NP}$-hard.*

Proof. To prove the computational complexity of the problem, we consider one of its special cases with a polynomial-time reduction from the 2-partition problem, which is shown to be $\mathcal{NP}$-hard (Garey and Johnson 1990). The 2-partition problem is defined as follows:

Let $A = \{a_1, a_2, ..., a_n\}$ be a finite set, where $a_i \in Z_{>0}$ is the size of element $i, i = 1, 2, ..., n$. Is there a subset $A' \subseteq A$ such that $\sum_{a_i \in A'} a_i = \sum_{a_i \in A - A'} a_i$?

We note that this problem remains $\mathcal{NP}$-hard even if $|A'| = |A|/2$ or if the elements in A are ordered (Garey and Johnson (1990).

We assume that the service provider offers two permit options 1 and 2, which are priced at $p_{i,1}$ and $p_{i,2}$, respectively, with equal total capacity such that $|M_1| = |M_2| = \sum_{a_i \in A} a_i/2$, and we then define the utility function of user $i, i = 1, 2, ..., n$ as follows:

$$
u_i(x) = \begin{cases} a_i + 1 & \text{if } x = a_{i,1} = a_{i,2} = a_i, \\ 0 & \text{otherwise.} \end{cases}
\tag{4.10}
$$

We assume that $p_{i,1} = p_{i,2} = a_i$ for $i = 1, 2, ..., n$. With this price menu and the utility functions defined above, each commuter $i, i = 0, 1, 2, ..., n$, presents a list which consists of both options. A solution for the above problem, if exists, will be a solution of the 2-partition problem and the objective function of the optimized model will be exactly equal to $\sum_{a_i \in A} a_i$. Since the above outlined reduction is a polynomial-time reduction from a 2-partition problem, we can now claim that UMPAP, too, is $\mathcal{NP}$-hard.

$\square$

4.2.2 A Heuristic for Mobility Permit Allocation

Under a socially acceptable equilibrium all users should be enabled to obtain favorable reservations closer to their most-desired time. Unfortunately, some conventional allocation schemes such as FCFS, queue-position, or priority-based would fail to do so. However, this issue can be tackled by a random assignment approach that treats the users equally at the time of allocation. The heuristic random priority (RP) allocation scheme proposed here is a way to accommodate reservation requests and address commuters' concerns about equity and fairness. The RP scheme works in two steps: In the first step, the system planner chooses at random an ordered list of users. In the second step each user $i, i = 1, 2, ..., N$ is allocated such that the assignment scheme goes through their preferences and allocates their most-preferred option, if it is available at the time of request. Algorithm (1) shows the pseudo-code for RP allocation scheme. This randomized allocation scheme addresses users' fairness expectation and then resolves the computational complexity of the problem. It also induces a non-manipulable property in the sense that the equilibrium is robust against market participants' beliefs about each other.

The proposed heuristic assignment scheme is widely used in practice and it comes with several other attractive properties as well. This mechanism is strategically simple as market participants do not need to bother themselves collecting information about others' actions or strategies. It helps policy makers to obtain information about the true preferences of participants. In addition, participants who lack the information are not disadvantaged relative to sophisticated participants (Azevedo and Budish 2017). This implicitly embeds another notion of fairness.

Proposition 4.2.1. *The traffic pattern outcome achieved by RP allocation scheme is Pareto efficient and strategy-proof.*

Proof. To show that the proposed scheme can achieve a Pareto efficient traffic pattern, we need to prove that any allocation outcome through PR is Pareto efficient. We note that the proposed scheme is a serial dictatorship as it generates an arbitrary list of users. Then starting from the user on top of the list each user is matched with their top choices that is available. It continues in this way down the list until the last user is allocated. As a serial dictatorship allocation scheme is proven to be Pareto efficient and strategy-proof (Smith 2003, Svensson 1999), and since every outcome of our mobility permit assignment corresponds to a serial dictatorship allocation mechanism, it readily proves the proposition.

Algorithm 1 Pseudo-code for Multi-item RP Allocation Scheme

1: *Initialize:*
2: $PM \leftarrow \{(tp_1, p_1), \ldots, (tp_{|K|}, p_{|K|})\}$ `Price Menu`
3: $M \leftarrow \{M_1, \ldots, M_{|K|}\}$ `List of Capacities`
4: $A \leftarrow \{a_1, \ldots, a_{|N|}\}$ `List of Agents`
5: $Q \leftarrow \{q_1, \ldots, q_{|N|}\}$ `Preference List`
6: $\Sigma \leftarrow$ *Random ordering of A*
7: **procedure** ALLOCATE
8:     **for** $\sigma \in \{1, \ldots, |\Sigma|\}$ **do**
9:         i^{th} `ordered user` $\leftarrow \sigma(i)$
10:         $AL_i \leftarrow 0$ **and** $k \leftarrow 1$
11:         **while** $AL_i = 0$ **and** $k \leq K$ **do**
12:             $\hat{q}_{i,l} \leftarrow \arg\max_{l \in \{q_{i,1}, \ldots, q_{i,L}\} = q_i \in Q} q_{i,l} \delta^i_{lk} \leq M_k$
13:             **if** $\hat{q}_{i,l} > 0$ **then**
14:                 $AL_i \leftarrow 1$
15:                 $k \leftarrow k + 1$
16:                 $M_k \leftarrow M_k - \hat{q}_{i,l}$ `Update Capacity`
17:                 $\hat{\boldsymbol{q}}_i \leftarrow \hat{q}_{i,l}$
18:                 $\hat{\boldsymbol{p}}_i \leftarrow p(q_{i,l})$
19:             **else** $k \leftarrow k + 1$
20:     **return** $\hat{\boldsymbol{q}}, \hat{\boldsymbol{p}}$

4.2.3 Pricing of Mobility Permits: A Progressive Auction Method

A mobility service yields an operative (near) optimal traffic pattern if each user can be assigned to the highest available option in their requested service list while no capacity limit is violated under the effective market price of permits. In other words, the mobility service provider needs to allocate mobility permits through a pricing mechanism which can attain the best equilibrium (price menu) from the point of view of the users. To this end, we propose a progressive pricing algorithm with two different pricing and allocation approaches. The first approach solves our UMPAP problem at each iteration of the auction and uses a flat pricing method that raises the permit price of all time slot permits at a constant rate. The iterative pricing and allocation terminates when the revenue outcome of the current iteration falls below that of the most recent iteration. The second approach for the pricing and allocation of permits is a hybrid method through which the price of over-demanded time slots are only raised at each step until there is no over-demanded time slot.

Figure (4.4) illustrates the general process and each step of the iterative hybrid method. Given each user's travel priorities, the mobility service provider offers an initial personalized price menu, a set of permitted trip options and the respective prices. As users place monetary value on each of the offered options, given a price menu, they can choose those items that bring non-negative surplus and puts the option that has the maximum surplus on top of their preference list and the second highest on the second position and so on. Then, users announces their preference list that specifies which permit option or options they want to buy at the initial prices. At each iteration and for each updated price menu, the service provider collects users' priorities, and for each user the mobility service provider offers a menu of available services which includes a permit vector and a price vector with one-to-one correspondence relation. The service provider then collects the users' preferences over the offered options and identifies a minimal over-demanded set, that is, an over-demanded set which none of its proper subsets is over-demanded (Demange et al. 1986). The service provider then raises the price of permits for the time slots in the minimal over-demanded set by one unit and announces a provisional allocation using our heuristic randomized allocation method. At each stage if the demand at each time slot falls shorter than the capacity for that slot then a market equilibrium has been already
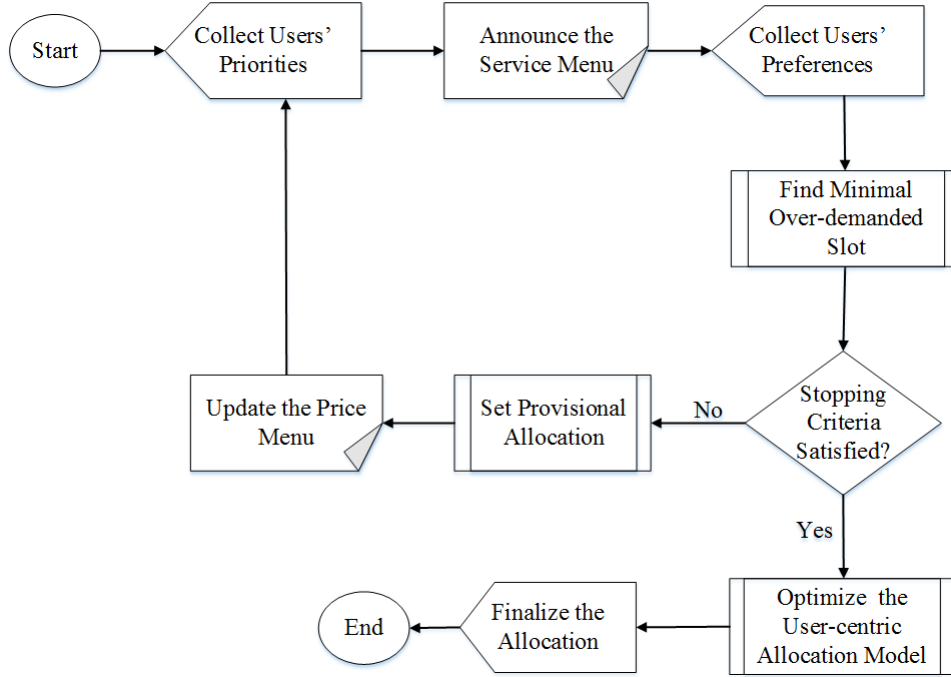
Figure 4.4 Iterative Auction for Pricing and Allocation of Permits

achieved for that time slot, i.e., it is possible to assign each user to one of their requested options. If no such assignment exists, then there is at least one over-demanded (congested) time slot that the number of users demanding permit for is greater than the capacity of that time slot. We modify the definition of an over-demanded set from Cramton et al. (2006). Let $q_i(P) = (q_{i,1}, q_{i,2}, ..., q_{i,L})$ be the preference list of user i at the price vector $P = (p_1, p_2, ..., p_L)$. Let $C \subseteq K$ and $R(C)$ be the set of users such that $0 \neq q_i(P) \subseteq C$ for all $i \in R(C)$. We say that a set of time slots C is over-demanded if $|R(C)| > \sum_{c \in C} M_c$. A necessary and sufficient condition for the existence of a feasible assignment is that there should be no over-demanded time slots due to Hall (1935) and Demange et al. (1986). According to Shapley and Shubik (1971), there exists an equilibrium, a unique price vector, that is the best equilibrium from the point of view of the users. This means that the iterative process should continue until no minimal over-demanded time slot exist. Now the planner performs the permit assignment process through optimizing the allocations according to our user-centric optimization model (UMPAP). In other words, when the decision is to stop the auction, the allocation is finalized based on the assignments generated through our user-centric mobility permit allocation model to find the optimum assignment outcome regarding the efficiency and equity concerns.

We note that the difference between the first and second auction methods relies on the way that provisional pricing and allocations and the termination criteria are performed, though both of them are similar regarding the other components. In the first auction method at iteration provisional allocations are done using the proposed UMPAP approach while in the second approach we replace it with the heuristic *RP* method to speed up the process. In addition, the second auction method uses the MODS pricing approach to raise the permit price of corresponding time slots while the first auction method simply raises the permit prices of all time slots with a constant ratio. The termination criteria in the first auction is based on comparing the revenue outcome of the current iteration with that of the most recent iteration. However, the second auction mechanism stops when no MODS exits, and it finalizes the permit assignments using the proposed optimization method (UMPAP) for permit allocation.

4.3 Performance Analysis

In this section, we use two different benchmark models under symmetric full information settings, developed in Appendix (A), for the purpose of evaluating the performance of the proposed schemes. The first benchmark model is efficiency-oriented mobility permit allocation under a coordinated system (*MPA-CS*) and system optimal (*MPA-SO*) setting. The *MPA-CS* problem models the pricing and allocation of permits as Stackelberg game with the service provider as the leader and the mobility users as followers. The *MPA-SO*, which is obtained from *MPA-CS*, only looks for optimum allocation of permits without pricing decision. The second benchmark model is mobility permit allocation under an equity-oriented system (*MPA-ES*). The *MPA-ES* problem reflects the Rawlsian notion of fairness principle under a symmetric case; it focuses on the individual utilities obtained by each user such that on the optimum solution even the worst off user can get the highest achievable utility (Rawls 2009).

4.3.1 Computational Experiments

Given different parameters in our scheme, we have considered a total of 16 problem sets with combinations of $N = 20$, and 40, and $M = 5$ and 10. We assume all the commuters have the same ideal time, t^* to pass the bottleneck. In all of the test problems we assume $L = 4$, i.e., the mobility service provider offers four time slots. In order to control adequately for heterogeneity, we randomly generate the parameters of the utility functions under two distinctive heterogeneity scenarios:

1) Proportional heterogeneity: where commuters' VOT, α, is an increasing function and follows the cumulative distribution function $F(x) = Pr\{\alpha \le x\}$ for $x \in [\underline{\alpha}, \bar{\alpha}]$. We denote $\bar{F}(x) = 1 - F(x)$ and the corresponding density function $f(\alpha)$, which is positive for $\alpha \in [\underline{\alpha}, \bar{\alpha}]$ and zero otherwise. For a specific commuter i, their early and late arrival penalty is proportional to α_i, i.e., $\beta = \rho_1 \alpha_i$ and $\gamma = \rho_2 \alpha_i$ where $0 < \rho_1 < 1 < \rho_2$ and ρ_1 and ρ_2 are identical for all commuters. This assumption on heterogeneity is similar to Vickrey (1973), Xiao et al. (2011) and Liu et al. (2015). To look at the efficiency loss due to user heterogeneity, we assume that $\alpha_i \sim U[\underline{\alpha}, \bar{\alpha}]$), i.e., the density function $f(\alpha) = 1/(\bar{\alpha} - \underline{\alpha})$, and $f(\alpha) = 0$ otherwise. In all our experiments we let $L = 4$, and $U_i = U\alpha_i$ where $U = 30$. Specifically, we assume $\bar{\alpha}/\underline{\alpha} = 2, 3, 4$, and 5. As summarized in Table (4.1), we consider 30 randomly generated problems for each of the Problem Sets 1-8, and $8 \times 30 = 240$ problems in total.

Table 4.1 Parameters Setup for Proportional Heterogeneity Setting

Problem Set	N	$M = N/s$	$\bar{\alpha}/\underline{\alpha}$	No. of Problems
1	20	5	2	30
2	20	5	3	30
3	20	5	4	30
4	20	5	5	30
5	40	10	2	30
6	40	10	3	30
7	40	10	4	30
8	40	10	5	30

For the first scenario the results of computational experiments are provided in Table (4.2). The first column "SO/Gr" shows the ratio of the total obtained utility in the centralized system (system optimal) to the utility of users in an ideal system where no bottleneck exists. The second column "CSPl/Gr" shows the ratio of the service provider's extracted utility, under a coordinated setting, to the utility of users in the ideal system. Similarly, the third column "CSUs/Gr" reports the ratio of the total welfare kept for the users under a coordinate setting. The sixth column "ES/Gr" reports the ratio of the preserved welfare under the equity-oriented system compared to the ideal system. Columns eight and nine, report the ratio of the extracted utility, under our MIP-based pricing and allocation scheme, to the total utility of the ideal system, respectively for the system planner "MIPPl/Gr" and the users of the system "MIPUS/Gr". In the same way, column 10 and 11 report the ratio of the extracted utility, under the hybrid pricing and allocation scheme, to the total utility from the ideal system, respectively for the system planner "HyPl/Gr" and the users of the system "HyUs/Gr". As it can be seen in the last row of Table (4.2), when the MIP-based pricing and allocation mechanism is applied on average both the system planner and users, under the privater information setting, can sustain around 48% of the total welfare from the ideal system. The efficiency loss is partly due to pricing and allocation decisions that the system planner applies when the system is not coordinated and partly because of the behavior of the assumed utility functions that restricts achieving the whole system's profit. However, our hybrid scheme under private information setting can achieve the same results that is achieved under the coordinated setting, under full information. With the proposed hybrid pricing and allocation scheme the system planner on average can extract more than 51% of the total welfare of the system while users can sustain about 46% of the welfare in the system, which are similar to the results obtained under the coordinated and full information setting.

Table 4.2 Results from Proportional Heterogeneity Setting

| Problem Set | MPA-SO | MPA-CS | | | MPA-ES | MIP Scheme (Flat Pricing) | | | Hybrid Scheme (MODS Pricing) | | |
	SO/Gr	CSPl/Gr	CSUs/Gr	Total	ES/Gr	MIPPl/Gr	MIPUS/Gr	Total	HyPl/Gr	HyUs/Gr	Total
1	97.83	68.90	28.93	97.83	97.44	66.09	31.41	97.50	69.44	28.77	98.21
2	98.03	54.79	43.24	98.03	97.39	52.94	44.48	97.42	55.19	43.01	98.20
3	98.11	46.16	51.95	98.11	97.39	43.94	53.51	97.45	46.48	51.77	98.25
4	98.18	40.63	57.55	98.18	97.41	38.06	59.43	97.48	40.87	57.48	98.35
5	97.84	67.06	30.78	97.84	97.68	64.17	33.31	97.48	66.51	30.60	97.11
6	98.02	52.42	45.61	98.02	97.86	50.21	47.29	97.50	52.25	45.41	97.66
7	98.09	43.52	54.57	98.09	97.94	41.49	56.01	97.49	43.27	54.43	97.70
8	98.20	37.17	61.03	98.20	98.07	34.69	62.81	97.50	37.32	60.96	98.29
Ave.	98.04	51.33	46.71	98.04	97.67	48.94	48.53	97.47	51.41	46.55	97.97

2) Non-proportional heterogeneity: is the case where commuters may have different ρ_1 and ρ_2. To focus on the heterogeneity in ρ_1 and ρ_2, we consider an identical VOT for all commuters. It is further assumed that $\eta = \rho_1/\rho_2$ is identical for all commuters. However, ρ_1 continuously increases from ρ_1^l to ρ_1^u with a cumulative distribution function of $F(x) = 1 - \bar{F}(x)$. Similarly, we denote $\bar{F}_1(x) = 1 - F_1(x)$. Note that, even though ρ_1 and ρ_2 can vary; it is assumed $0 < \rho_1 < 1 < \rho_2$; then we see that, $\rho_1^l < \rho_1^u < 1$ and $\rho_2^l = \eta\rho_1^l > 1$. This type of user heterogeneity is similar to that considered in Tian et al. (2013) and Liu et al. (2015) and many others. Specifically, in this scenario we assume ρ_1 is uniformly distributed as $\rho_1 \sim U(0.5, 1)$. Similarly, we let $L = 4$, and $U_i = 60$. As summarized in Table (4.3), for this scenario, we consider 30 randomly generated problems for each of the Problem Sets 1-8, and $8 \times 30 = 240$ problems in total.

Table 4.3 Parameters Setup for Non-proportional Heterogeneity

Problem Set	N	$M = N/s$	$\eta = \rho_1/\rho_2$	No. of Problems
1	20	5	2	30
2	20	5	3	30
3	20	5	4	30
4	20	5	5	30
5	40	10	2	30
6	40	10	3	30
7	40	10	4	30
8	40	10	5	30

For the second scenario the results of computational experiments are provided in Table (4.4) with the same representative columns as described in Table (4.2).

Table 4.4 Results from Non-proportional Heterogeneity Setting

	MPA-SO	*MPA-CS*			*MPA-ES*	MIP Scheme (Flat Pricing)			Hybrid Scheme (MODS Pricing)		
Problem Set	SO/Gr	CSPl/Gr	CSUs/Gr	Total	ES/Gr	MIPPl/Gr	MIPUS/Gr	Total	HyPl/Gr	HyUs/Gr	Total
1	91.32	88.34	2.98	91.32	91.17	79.57	4.54	84.11	87.10	4.34	91.44
2	90.03	86.85	3.18	90.03	89.70	77.31	5.57	82.88	83.33	6.45	89.78
3	88.44	84.95	3.49	88.44	87.85	72.10	8.84	80.94	83.31	4.91	88.22
4	86.76	83.08	3.67	86.75	86.01	65.59	13.15	78.74	78.92	5.37	84.30
5	91.49	88.25	3.24	91.49	91.45	80.30	3.43	83.73	86.86	4.71	91.57
6	89.95	86.47	3.48	89.95	89.75	77.43	5.16	82.59	84.98	4.58	89.56
7	88.57	85.00	3.58	88.57	88.43	72.32	7.98	80.30	83.34	5.08	88.42
8	87.32	83.41	3.91	87.32	86.52	66.73	11.89	78.62	81.50	5.48	86.98
Ave	89.44	85.79	3.44	89.23	88.86	73.92	7.57	81.49	83.67	5.12	88.78

As it can be seen in the last row of Table (4.4), with the MIP-based pricing and alloca-tion mechanism, on average the system planner can sustain around 74% of the total welfare of the ideal system and the users on average can get 7% of the total welfare. However, our hybrid scheme, under private information setting, can achieve efficiency results similar to that under the coordinated setting, under full information. With our hybrid permit pricing and allocation scheme the service provider on average can extract more than 83% of the total welfare of the system while users can sustain about 5% of the welfare in the system, which are very close to the results obtained under the coordinated setting. The efficiency loss of MIP-based method can be attributed to flat pricing and allocation decisions. While with the hybrid scheme the efficiency loss is less than 1% which is trivial. The overall comparison of the results between these two scenario show that user heterogeneity causes further efficiency loss, though the proposed auction mechanisms can mitigate a propor-tion of the loss. We can see that under non-proportional heterogeneity setting the hybrid auction can transfer more surplus from users to the service provider.

From Tables (4.2) and (4.4), one can also see that the equity-oriented mobility permit allocation scheme can sustain about 89% to 97% of the overall welfare in the system. In other words, 3% to 11% of total welfare is lost when equity is the sole objective within the system having a bottleneck. The welfare lost with MIP-based scheme can reach to 16% of the total welfare within the system, mostly due to the flat pricing scheme. However, the hybrid scheme with the MODS pricing method loses a trivial proportion (less than 0.2%) of total welfare within the mobility system under both proportional and non-proportional heterogeneity settings.

We now turn into analyzing the overall performance and acceptability ratios of the proposed schemes (MIP-based and Hybrid) in terms of efficiency and equity loss using POF and equity POE metrics defined in Section (3.2.6). Table (4.5) presents the values of PoF and PoE for proportional and non-proportional heterogeneity settings for both MIP-based and hybrid mobility permit pricing and allocation schemes. The values of PoF show that the loss of total utility or overall welfare of a fair solution compared to the system optimum (MPA-SO) in the non-proportional heterogeneity setting is significantly less than those in the proportional heterogeneity setting. Moreover, the hybrid pricing and allocation scheme results in a trivial efficiency losses, PoF is about 17% under the proportional heterogeneity setting and 6% under the non-proportional heterogeneity setting. This difference can be attributed to less diversity of users in terms of the earliness and lateness under non-proportional heterogeneity setting. We see that the value of PoE for both schemes (MIP-

70

based and Hybrid) are very close to each other and trivial under both settings while with the hybrid mobility pricing and allocation scheme almost no efficiency is lost compared to the overall welfare under system optimum traffic pattern.

Table 4.5 The values of *PoF* and *PoE* for Proportional and Non-proportional Heterogeneity Settings

	Pro. Hete.		Non-pro. Hete.	
Criteria	MIP Scheme	Hybrid Scheme	MIP Scheme	Hybrid Scheme
PoF	0.50	0.48	0.17	0.06
PoE	0.02	0.00	0.08	0.00

The computational effort required to perform pricing and allocation of mobility permits is highly dependent to the number of participating users and the number of available options. On average, the computational effort grows as the number of users increases. However, to provide a rough estimate in the largest and complex problems (i.e., Problem Set 4-8 of non-proportional heterogeneity scenario), the proposed MIP-based scheme on average can finalize the pricing and allocation decisions in around 1450 s while with the proposed hybrid mechanism on average it takes about 840 s. We note that the most computational time in the hybrid scheme is due to the final step that requires the permit allocation problem to be solved optimally. To solve this issue, a heuristic allocation scheme can be used to finalize permit allocations when the respective market prices are found through the hybrid scheme. This can significantly reduce the computational time, however, it may result in transferring a small proportion of the service provider's utility to mobility users.

4.4 Conclusion

In this chapter, we focused on the design and analysis of user-centric pricing and allocation of mobility permits for roadways with one bottleneck. We dealt with observing operational objectives, particularly, balancing efficiency and fairness in mobility permit allocation. We then explored the theoretical properties of the proposed scheme and showed that the proposed scheme can achieve an optimal traffic pattern; though, it is computationally intensive to solve large size problem instances. Next, to tack the computational

complexity of the proposed scheme, we proposed a heuristic permit allocation algorithm that sustains a Pareto-optimal traffic pattern with less computational effort. Next, we designed the steps of an iterative auction mechanism for pricing the mobility permit under two different pricing methods. To analyze the performance of the proposed schemes, we performed computational experiments under different parameter settings. We showed that a hybrid mechanism with a minimal over-demanded set pricing and heuristic allocation method can be a good candidate for being the mobility scheme component of the proposed integrated user-centric traffic management system. Although, the presented MP-based traffic management scheme is Pareto-improving and most probable to be economically and socially acceptable, to completely evaluate its acceptability we need to take into account other practical factors, such as the way in which the revenue outcome of the scheme is used. This needs designing revenue recycling policies that enhance the acceptability of mobility management reforms (Mayeres and Proost 2002), which is out of this research's scopes and objectives.

Chapter 5

A Mobility Credit Scheme for Managing Traffic on Transportation Networks with Advanced Connectivity Technologies [†]

5.1 Summary

In this chapter, we present a credit-based mobility scheme to regulate traffic flow on a road network consisting of a mix of CAVs and non-CAVs. Particularly, a set of revenue-neutral credit schemes are proposed that uses a combination of link-based rewards and charges to actuate network-wide travel behavior towards a system optimum pattern. We first formulate the underlying problem as a logit-based stochastic traffic assignment problem, using a path-free mathematical program. We then model the mixed vehicle stochastic user- and system-optimal traffic assignment problems as non-linear complementarity problems (NCPs) and use them to find Pareto-improving link-specific charges and rewards. Numerical analyses demonstrate the efficiency of the proposed scheme under different hypothetical scenarios and the performance of different variants of the proposed mobility

---

scheme. The study results can be used to support the effectiveness of credit-based scheme in leveraging the advantage of CAVs in moving towards the marketable mobility paradigm.

5.2 A Compound Credit-based Scheme

5.2.1 User Equilibrium

We consider a transportation network represented by a fully-connected directed graph $G(N, A)$ with multiple OD pairs. Each node is denoted by a natural number i, and each link $a = (i, j)$ of A is identified by the upstream node i and the downstream node j. The node set N includes origin nodes $o \in O$ from which users start their trips, and destination nodes $d \in D$ at which users terminate their trips. Let W denote the set of OD pairs, R denote the set of all routes, and $R_w \subset R$ be the set of all routes between an OD pair $w \in W$. The travel demand for each OD pair $w = (o, d) \in W$ is denoted by q_w, $q_w > 0$ and the total travel demand is given by $q = \sum_{w \in W} q_w$.

For simplicity, let $t_a(v_a)$, $\forall a \in A$ denote the travel time (performance function) for link a; therefore, the travel time on route $r \in R_w$ can be stated as $t_r^w = \sum_{a \in A} t_a(v_a)\delta_{ar}^w, r \in R_w, w \in W$, where δ_{ar}^w is equal to 1 if path $r \in R_w$, $w \in W$ uses link $a \in A$; and zero otherwise.

A traffic agency (central transport authority) wishes to use a credit-based charging scheme denoted by (K, κ) with which he circulates K number of credits in the mobility market. Travelers are either charged with $\kappa_a > 0$ or subsided with $\kappa_a < 0$ if they traverse on link a; otherwise, it will be for free, i.e., $\kappa_a = 0$. The mobility market consists of two different classes of users who share the same network, namely, cooperative and non-cooperative users. The cooperative users, CAVs, are assumed to follow a network-wide routing principle, making more informed real-time route choice decisions. On the other hand, the non-cooperative users, non-CAVs, adopt self-serving routing principle to minimize their individual travel times. We note that all users must have registered with a central transport authority in order for the scheme to work. Expectedly, tracking technologies along with IoT will create a unified gateway to achieve this.

We assume that the traffic management agency uses a compound scheme under which non-CAVs and CAVs users can be subject to different charge and subsidy policies. This

means the charging are differentiated based on their registered class. In all these cases the charges and rewards would be based on the cooperation level of the active vehicle on following spontaneous network-wide routing options. In other words, changing the driver mode to drive a CAV in human mode but cooperative manner is allowed as far as the user follow the centrally made routing decisions. However, if the users opt out of the cooperative behavior then their cost would be based on their self-serving routing decisions. We need to distinguish the CAV users who make informed routing decisions, i.e., a CAV on human mode does not change anything as far as it follows informed centrally made routing decisions. Another point that we need to mention is that our proposed credit scheme is a unified scheme where the two types of users could trade the credits, however, it is the charge and reward rates on the links that differentiate the scheme for them. Moreover, there is no initial allocation impact due to the credit scheme. It is assumed that the integration of credit scheme into ATIS/CAVs technologies would lead to negligible (zero) transaction costs, which makes the final equilibrium to be independent of initial allocations of the credits (Yang and Wang 2011). Essentially, upon following the spontaneous system-wide provided routing options, users are either charged or rewarded depending on their cooperation level.

A differentiated scheme is suitable for the heterogeneous situations where undifferentiated charges might not be sufficient for achieving a system optimal state (Mehr and Horowitz 2019). Another benefit of having different charge and reward rates is that it enables the system planner to promote travel behavior of the corresponding participants towards a common goal (e.g., maximizing social welfare, promoting desired traffic condition or an emission control program). Paying one-shot purchase tax or simply applying tolls, through having some motivational justification, would not help promote an improved travel decision behavior. Indeed, reward-based instruments, compared to single-shot subsidy or purchase tax schemes, are shown to be more effective in promoting eco-driving travel decision behavior (Tsirimpa et al. 2019). We assume that a network-wide monitoring infrastructure is in place in which participants can obtain real-time travel decision aids and users' movements throughout the network can be monitored. Thus the participants can be charged or rewarded depending on their cooperation level tied with their travel decision behaviors. The credit-based scheme aims to stimulate a network-wide optimum state by balancing the spent travel time (cost) with the incurred monetary charges or rewards such that those who look for short travel time based on UE routing principle must pay for it, and those who obey cooperative eco-driving behavior based on SO routing principle are

rewarded accordingly.

For these settings the effects of stochastic behavior and the level of cooperation can be incorporated by different values of the variability parameters (Yang 1998, Yin and Yang 2003). Without loss of generality, we assume two different levels of operational uncertainty for both type of users. A higher variability level is attributed to non-CAVs users mainly due to the fact that those users follow the UE routing principle to minimize their individual travel times. A lower level of variability is attributed to CAV users to account for the fact that they share common infrastructures such as road links with non-CAVs, and are subject to network interruption and other possible situations such as bad weather, moving bottlenecks, delays and lags affecting vehicle reaction and response times.

With these explanations, the first class of travelers are non-CAV drivers, who rest on a user-optimum manner to route themselves under a stochastic setting, accounting $\hat{q} < q$ part of the total travel demand. It is further assumed that the routing decisions of non-CAV drivers are associated with a variability parameter $\hat{\theta}$ and a charging scheme $\hat{\kappa} \in \kappa$. The variability parameter is to include the monetary equivalent of the cost associated with class-specific travel behavior and vehicle features. The non-CAVs make their route choice following the SUE principle and select the route r from the route set $r \in R_w$ based on minimal perceived travel costs $(\hat{c}_r^w)$ defined as follows:

$$\hat{c}_r^w = \sum_{a \in A}(t_a(v_a) + \rho\hat{\kappa}_a)\delta_{ar}^w + \hat{\epsilon}_r^w. \tag{5.1}$$

where ρ denotes the monetary equivalent of unit credit price in the credit market.

With this definition, the perceived travel cost becomes equal to the monetary equivalent of generalized travel time on the path $r \in R_w$ which is the sum of the path travel time and total credit used/earned on that path plus users' perception error $(\hat{\epsilon}_r^w)$ on the path travel time, which follows Gumbel distribution for logit-based SUE. More specifically, we assume that this class of users choose the rth route with probability $\hat{p}_r^w$ which reads by a logit-based formula:

$$\hat{p}_r^w \quad = \quad \frac{exp(\frac{\hat{c}_r^w}{-\hat{\theta}})}{\sum_{k \in R_w} exp(\frac{\hat{c}_k^w}{-\hat{\theta}})} \tag{5.2}$$

$$= \quad \frac{exp\left[\frac{t_r^w + \rho \sum_{a \in A} \hat{\kappa}_a \delta_{a,r}}{-\hat{\theta}}\right]}{\sum_{k \in R_w} exp\left[\frac{t_k^w + \rho \sum_{a \in A} \hat{\kappa}_a \delta_{a,r}}{-\hat{\theta}}\right]}, r \in R_w, w \in W, \tag{5.3}$$

where $\hat{\theta}$ is travelers' perception dispersion parameter, i.e., a positive value related to the standard deviation of the random term, and measures the sensitivity of route choices to travel cost.

Therefore, the traffic assignment problem of non-CAV drivers would satisfy the following traffic pattern equilibrium:

$$\hat{f}_r^w = \hat{p}_r^w \hat{q}_w, r \in R_w, w \in W, \tag{5.4}$$

where $\hat{q}_w$ is the demand of non-CAV users between OD pair $w \in W$, such that $\sum_{w \in W} \hat{q}_w = \hat{q}$, and $\hat{f}_r^w$ is the path flow of these users on path $r \in R_w$.

The second class of travelers are CAV users who constitute $\bar{q} = q - \hat{q}$ part of the total travel demand. It is assumed that CAV users obtain partial information about the road traffic condition; therefore, they are associated with a different variability parameter $\bar{\theta}$ and a different charging scheme $\bar{\kappa} \in \kappa$. Therefore, their experienced generalized travel time is $\bar{c}_r^w = \sum_{a \in A}(t_a(v_a) + \rho \bar{\kappa}_a)\delta_{ar}^w + \bar{\epsilon}_r^w$ on the path $r \in R_w$, and their traffic assignment would satisfy the following traffic pattern:

$$\bar{f}_r^w = \bar{p}_r^w \bar{q}_w, r \in R_w, w \in W, \tag{5.5}$$

where $\bar{q}_w$ denotes the demand of CAV users between OD pair $w \in W$ such that $\sum_{w \in W} \bar{q}_w = \bar{q}$.

It is common in the literature that when generalized costs are created on the basis of travel time and money, features such as higher average speed (travel efficiency) and fuel consumption costs are converted into equivalent monetary costs. All these factors can

be seen in an extended form of generalized travel cost function. For example, to study the effects of CAV ownership on trip, mode, and route choice, Levin and Boyles (2015) use an extended form of generalized travel time function with separate factors attributed to each different feature. Similarly, to account for the reduced energy consumption due to platooning, Wang, Peeta and He (2019) use a discount factor for fuel consumption parameter in link travel time. In this study, we keep the travel time in the simplistic form, but differentiate the generalized costs of CAVs and non-CAVs using different charging scheme and associating different variability parameters to their travel time, with a lower variability representing a higher travel efficiency of CAVs.

Under a feasible compound credit scheme $(K, \hat{\kappa}, \bar{\kappa})$, the mixed-vehicle stochastic user equilibrium (MSUE) traffic flow and market equilibrium (ME) conditions can be written as:

$$\hat{f}_r^w = \hat{p}_r^w \hat{q}_w, \ \ r \in R_w, w \in W, \tag{5.6}$$

$$\bar{f}_r^w = \bar{p}_r^w \bar{q}_w, \ \ r \in R_w, w \in W, \tag{5.7}$$

$$(K - \sum_a (\hat{\kappa}_a \hat{v}_a + \bar{\kappa}_a \bar{v}_a))\rho = 0, \rho \geq 0, \sum_a \hat{\kappa}_a \hat{v}_a + \bar{\kappa}_a \bar{v}_a \leq K, \tag{5.8}$$

where

$$\hat{v}_a = \sum_{w \in W} \sum_{r \in R^w} (\hat{f}_r^w) \delta_{ar}^w,$$

$$\bar{v}_a = \sum_{w \in W} \sum_{r \in R^w} (\bar{f}_r^w) \delta_{ar}^w, a \in A. \tag{5.9}$$

Equations (5.6) to (5.8) represent mixed stochastic user equilibrium and the credit market clearing conditions. Specifically, Equations (5.6) and (5.7) is the traffic flow distribution based on SUE, and Equation (5.8) states that the equilibrium credit scheme $(K, \hat{\kappa}, \bar{\kappa})$ is effective ($\rho \neq 0$) only when the obtained link flow pattern under the compound credit scheme is not identical to UE mixed link flow pattern. To assure this would not be the case, we assume $\sum_a (\hat{\kappa}_a \hat{v}_a^0 + \bar{\kappa}_a \bar{v}_a^0) > K$ for the mixed link flow pattern $(\hat{v}_a^0, \bar{v}_a^0)$ under no-credit scheme intervention. A similar argument can be found in Han and Cheng (2017) for the single class SUE network flow pattern.

When we set $K = 0$, we will obtain a revenue-neutral cyclic charging scheme which is different from periodic credit distribution scheme. In the cyclic scheme credits never expire but circulate within the system such that the total number of collected credits is constrained to be zero (Xiao, Long, Li, Kou and Nie 2019). Moreover, the cyclic credit charging scheme allows for negative link charges (or subsidies), therefore users may obtain some subsidy from choosing links with negative charges. This means that a cyclic credit charging scheme under an ATIS has more flexibility than a periodic credit distribution scheme to incentivize users to observe centralized routing decisions. A compound scheme implicitly can include an undifferentiated case where the traffic agency uses the same charging scheme for all vehicles just by setting $\hat{\kappa}_a = \bar{\kappa}_a, \forall a \in A$.

5.2.2 Equivalent Mathematical Programming Formulation

Maher et al. (2005) propose a closed-form expression for the stochastic traffic assignment with Gumbel distributed perception error. Sharing the same network, our setting is different as the two classes of network users would interact and impact each other which requires a different modeling approach. In this section, we propose the following logit-based minimization program model for characterizing the mixed traffic behavior and the resultant equilibrium with two different Gumbel distributed perception errors under a feasible compound credit-based travel demand scheme $(K, \hat{\kappa}, \bar{\kappa})$.

$$Min\ Z(\hat{\boldsymbol{f}}, \bar{\boldsymbol{f}}) = \sum_{a \in A} \int_0^{\hat{v}_a + \bar{v}_a} t_a(x)\,dx$$

$$+ \ \hat{\theta} \sum_{w \in W} \sum_{r \in R^w} \hat{f}_r^w (\ln \hat{f}_r^w - 1) + \bar{\theta} \sum_{w \in W} \sum_{r \in R^w} \bar{f}_r^w (\ln \bar{f}_r^w - 1)$$

$$- \ \hat{\theta} \sum_{w \in W} \hat{q}_w (\ln \hat{q}_w - 1) - \bar{\theta} \sum_{w \in W} \bar{q}_w (\ln \bar{q}_w - 1) \tag{5.10}$$

subject to:

$$\sum_{r \in R^w} \bar{f}_r^w = \bar{q}_w, w \in W, \tag{5.11}$$

$$\sum_{r \in R^w} \hat{f}_r^w = \hat{q}_w, w \in W, \tag{5.12}$$

$$\sum_{a \in A} \hat{\kappa}_a \hat{v}_a + \bar{\kappa}_a \bar{v}_a \leq K, \tag{5.13}$$

$$\bar{\kappa}_a \leq \hat{\kappa}_a, a \in A, \tag{5.14}$$

$$\bar{f}_r^w \geq 0, \ r \in R^w, w \in W, \tag{5.15}$$

$$\hat{f}_r^w \geq 0, \ r \in R^w, w \in W, \tag{5.16}$$

where the link flows are defined by

$$\hat{v}_a = \sum_{w \in W} \sum_{r \in R^w} (\hat{f}_r^w) \delta_{ar}^w, a \in A, \tag{5.17}$$

$$\bar{v}_a = \sum_{w \in W} \sum_{r \in R^w} (\bar{f}_r^w) \delta_{ar}^w, a \in A. \tag{5.18}$$

The above simplified model looks like Fisk's logit-type SUE model with two different perception errors $(\hat{\theta}, \bar{\theta})$, restrained by additional credit feasibility constraints (5.13) and (5.14). Specifically, the constraint set (5.14) guarantees a permissive charging scheme to encourage users to switch to CAVs driving mode. However, for simplicity and without loss of generality, we relax this constraint here. In Section (5.2.5), we will show how we can incorporate it when designing a compound credit charging scheme. Moreover, we will also present a minimal value charging scheme such that the absolute value of charges, i.e., $|\bar{\kappa}_a|$ and $|\hat{\kappa}_a|$ for all $a \in A$, are expected to be as low as possible to address the concerns pertinent to social equity by obtaining small values of charges or rewards which are distributed over the entire network. This is elaborated in Section (5.2.5).

We note that the last two terms in the objective function (5.10) can be omitted, as they are constant. The proposed mathematical model can be reduced to the case where CAV users are assumed to have perfect compliance ($\bar{\theta} = 0$) and thus make link choice decisions in a system-optimum manner and non-CAV users follow UE routing principle under stochastic setting. We assume that, when driving on each link, users stick to their most recently chosen travel behavior before entering the link. Therefore aggregate number of users of each type would not change due to spontaneous class type switches. This situation can also be modeled by replacing the objective function (5.10) with the following objective functions.

$$
\begin{aligned}
Z(\hat{\boldsymbol{f}}, \bar{\boldsymbol{f}}) \;=\; & \sum_{a \in A} \bar{v}_a t_a(\hat{v}_a + \bar{v}_a) + \\
& \sum_{a \in A} \int_0^{\hat{v}_a + \bar{v}_a} t_a(x)\,dx + \hat{\theta} \sum_{w \in W} \sum_{r \in R^w} \hat{f}_r^w (\ln \hat{f}_r^w - 1)
\end{aligned}
\tag{5.19}
$$

It can also be reduced to a deterministic traffic stream situation where non-CAVs user with user-equilibrium behavioral pattern are mixed with CAVs users with system-optimal behavioral pattern by replacing the objective function (5.10) with the following objective function.

$$
Z(\hat{\boldsymbol{f}}, \bar{\boldsymbol{f}}) \;=\; \sum_{a \in A} \int_0^{\hat{v}_a + \bar{v}_a} t_a(x)\,dx + \sum_{a \in A} \bar{v}_a t_a(\hat{v}_a + \bar{v}_a)
\tag{5.20}
$$

In both of these situations the CAV users would be free of charge, however their travel decisions can be managed centrally to control and distribute the traffic demand of non-CAV users. The latter scenario with adjustments will be an extension of Chen et al. (2020) to a stochastic traffic flow setting. In this chapter, we keep our formulation in its general form modeled in the minimization programming problem (5.10)-(5.16). However, the obtained results are relevant with corresponding adjustments to the described scenarios as well.

Proposition 5.2.1. *(Existence) For a convex and monotonically increasing (with the amount of flow) link performance functions, the minimization problem (5.10) - (5.18) is a convex programming problem with linear constraints and has a unique aggregate link ow solution for each vehicle class.*

Proof. We substitute Equations (5.17) and (5.18) into the objective function (5.10) such that it can be stated only in terms of path flow variables $(\hat{\boldsymbol{f}}, \bar{\boldsymbol{f}})$.

$$
\begin{aligned}
Z(\hat{\boldsymbol{f}}, \bar{\boldsymbol{f}}) &= \sum_{a \in A} \int_0^{\sum_{w \in W} \sum_{r \in R^w} (\hat{f}_r^w + \bar{f}_r^w) \delta_{ar}^w} t_a(x) \, dx \\
&+ \hat{\theta} \sum_{w \in W} \sum_{r \in R^w} \hat{f}_r^w (\ln \hat{f}_r^w - 1) + \bar{\theta} \sum_{w \in W} \sum_{r \in R^w} \bar{f}_r^w (\ln \bar{f}_r^w - 1).
\end{aligned}
\tag{5.21}
$$

We then need to determine its Hessian matrix that is

$$
\boldsymbol{H}[Z(\hat{\boldsymbol{f}}, \bar{\boldsymbol{f}})] = \begin{bmatrix} \frac{\partial^2 Z(\hat{\boldsymbol{f}}, \bar{\boldsymbol{f}})}{\partial \hat{f}_r^w \partial \hat{f}_{r'}^{w'}} & \frac{\partial^2 Z(\hat{\boldsymbol{f}}, \bar{\boldsymbol{f}})}{\partial \hat{f}_r^w \partial \bar{f}_{r'}^{w'}} \\ \frac{\partial^2 Z(\hat{\boldsymbol{f}}, \bar{\boldsymbol{f}})}{\partial \bar{f}_r^w \partial \hat{f}_{r'}^{w'}} & \frac{\partial^2 Z(\hat{\boldsymbol{f}}, \bar{\boldsymbol{f}})}{\partial \bar{f}_r^w \partial \bar{f}_{r'}^{w'}} \end{bmatrix}.
\tag{5.22}
$$

The dimension of this matrix is $(RW)^4$. We drive the first and second order partials of $Z(\hat{\boldsymbol{f}}, \bar{\boldsymbol{f}})$ with respect to different possible combinations of path flow variables for non-CAVs and CAVs. Its first derivative with respect to a given non-CAVs' path flow variable, $(\hat{f}_r^w)$, is

$$
\frac{\partial Z(\hat{\boldsymbol{f}}, \bar{\boldsymbol{f}})}{\partial \hat{f}_r^w} = \sum_{a \in A} t_a \Big(\sum_{w \in W} \sum_{r \in R^w} (\hat{f}_r^w + \bar{f}_r^w) \delta_{ar}^w \Big) \delta_{ar}^w + \hat{\theta} \ln \hat{f}_r^w = t_r^w + \hat{\theta} \ln \hat{f}_r^w.
\tag{5.23}
$$

Where t_r^w is the travel time on path r between OD pair $w \in W$, that is, $t_r^w = \sum_{a \in A} t_a(\sum_{w \in W} \sum_{r \in R^w} (\hat{f}_r^w + \bar{f}_r^w) \delta_{ar}^w) \delta_{ar}^w$. In the same spirit, the second derivative of objective function (5.21) with respect to a second path flow variable of non-CAVs becomes

$$
\frac{\partial^2 Z(\hat{\boldsymbol{f}}, \bar{\boldsymbol{f}})}{\partial \hat{f}_r^w \partial \hat{f}_{r'}^{w'}} = \begin{cases} \partial(t_r^w)/\partial \hat{f}_r^w + \frac{\hat{\theta}}{\hat{f}_r^w} & \text{if } r' = r; w' = w, \\ \partial(t_r^w)/\partial \hat{f}_{r'}^{w'} & \text{otherwise.} \end{cases}
\tag{5.24}
$$

Also we can drive the corresponding second derivative of the objective function (5.21) with respect to two path flow variables of CAVs' as follows:

$$\frac{\partial^2 Z(\hat{\boldsymbol{f}}, \bar{\boldsymbol{f}})}{\partial \bar{f}_r^w \partial \bar{f}_{r'}^{w'}} = \begin{cases} \partial(t_r^w)/\partial \bar{f}_r^w + \frac{\bar{\theta}}{\bar{f}_r^w} & \text{if } r' = r; w' = w, \\ \partial(t_r^w)/\partial \bar{f}_{r'}^{w'} & \text{otherwise.} \end{cases} \qquad (5.25)$$

Moreover, we can drive the corresponding derivative with respect to both non-CAVs' and CAVs' path flow variables, which becomes

$$\frac{\partial^2 Z(\hat{\boldsymbol{f}}, \bar{\boldsymbol{f}})}{\partial \hat{f}_r^w \partial \bar{f}_{r'}^{w'}} = \begin{cases} \partial(t_r^w)/\partial \bar{f}_r^w & \text{if } r' = r; w' = w, \\ \partial(t_r^w)/\partial \bar{f}_{r'}^{w'} & \text{otherwise.} \end{cases} \qquad (5.26)$$

We have $\hat{f}_r^w > 0$ and $\bar{f}_r^w > 0$ for $r \in R_w, w \in W$, based on our assumptions, and that the path travel times (t_r^w) are sum of the corresponding link performance functions that are monotonically increasing with respect to their flow and that the variability parameters $\hat{\theta}$ and $\bar{\theta}$ are positive values. Considering Equations (5.24)-(5.26), and the fact that links performance functions are concave and monotonically increasing with respect to their corresponding flows, we have positive definite matrices $(\partial(t_r^w)/\partial \bar{f}_r^w)$ and two positive definite matrices $(\frac{\hat{\theta}}{\hat{f}_r^w}, \text{ and } \frac{\bar{\theta}}{\bar{f}_r^w})$. Overall, we can see that the Hessian matrix is the summation of positive definite and positive definite matrices, and thus it is a positive definite matrix. Therefore, the objective function of minimization programming is strictly convex with respect to path flows. Also, other constraints as well as the credit feasibility constraints are linear, thus the problem is a convex programming with linear constraints. Moreover, the uniqueness of solution route and link flows can be guaranteed by the strict convexity of the objective function and solution space (Yang 1998). This completes the proof. □

This proposition means that for monotone increasing link performance functions, the equilibrium link flow pattern $(v_a = \hat{v}_a + \bar{v}_a)$ under the compound credit scheme $(K, \hat{\kappa}, \bar{\kappa})$, if exists, is unique.

Proposition 5.2.2. *(Equivalence) Any local minimum solution $(\bar{\boldsymbol{f}}^*, \hat{\boldsymbol{f}}^*)$ of the optimization model satisfies the conditions (5.6)-(5.8), and the Lagrangian multiplier associated with the linear credit feasibility constraint is the market equilibrium credit price.*

Proof. Denote the Lagrangian function of the minimization problem (5.10)-(5.16) as below

$$L(\bar{\bm{f}}, \hat{\bm{f}}, \lambda, \mu, \rho) = Z(\bar{\bm{f}}, \hat{\bm{f}}) + \rho(K - \sum_a (\hat{\kappa}_a \hat{v}_a + \bar{\kappa}_a \bar{v}_a))$$

$$+ \sum_{w \in W} \lambda_w \{ \sum_{r \in R_w} \bar{f}_r^w - \bar{q}_w \} + \sum_{w \in W} \mu_w \{ \sum_{r \in R_w} \hat{f}_r^w - \hat{q}_w \}, \qquad (5.27)$$

where λ_w, μ_w, and ρ are the Lagrange multipliers associated with constraints (5.11), (5.12) and (5.13), respectively. The Lagrange multipliers λ_w and μ_w can be regarded as the minimal generalized travel costs, respectively, for CAVs and non-CAVs. In the same way, the Lagrange multiplier ρ can be regarded as the credit price.

From the Karush-Kuhn-Tucker (KKT) conditions of Lagrangian function (5.27), the following necessary and sufficient conditions can be obtained,

$$\partial \frac{L(\bar{\bm{f}}, \hat{\bm{f}}, \lambda, \mu, \rho)}{\partial \hat{\bm{f}}} \hat{f}_r^w = 0, \qquad (5.28)$$

$$\partial \frac{L(\bar{\bm{f}}, \hat{\bm{f}}, \lambda, \mu, \rho)}{\partial \hat{\bm{f}}} = 0, \hat{f}_r^w \geq 0, \qquad (5.29)$$

$$\partial \frac{L(\bar{\bm{f}}, \hat{\bm{f}}, \lambda, \mu, \rho)}{\partial \bar{\bm{f}}} \bar{f}_r^w = 0, \qquad (5.30)$$

$$\partial \frac{L(\bar{\bm{f}}, \hat{\bm{f}}, \lambda, \mu, \rho)}{\partial \bar{\bm{f}}} = 0, \bar{f}_r^w \geq 0, \qquad (5.31)$$

$$(K - \sum_a (\hat{\kappa}_a \hat{v}_a + \bar{\kappa}_a \bar{v}_a))\rho = 0, \qquad (5.32)$$

$$K - \sum_a (\hat{\kappa}_a \hat{v}_a + \bar{\kappa}_a \bar{v}_a) \geq 0, \ \ \rho \geq 0. \qquad (5.33)$$

First we consider the route flow associated with the non-CAV drivers. Equating the

partial derivatives (5.29) to zero we can obtain the following optimality conditions:

$$\hat{\theta} \ln \hat{f}_r^w + \sum_{a \in A} (t_a(v_a) + \rho \hat{\kappa}_a) \delta_{ar}^w - \mu_w = 0, r \in R_w, w \in W, \tag{5.34}$$

We note that this equation is only valid when $\hat{f}_r^w > 0$, for $r \in R_w, w \in W$, since, all path with zero flow are inactive and essentially have a generalized travel cost larger than those of active path for the same OD pairs, due to UE conditions.

Now, we focus on the routes belonging to a specific OD pair $w \in W$, denoting $t_r^w(\bullet) = \sum_{a \in A} (t_a(v_a) + \rho \hat{\kappa}_a) \delta_{ar}^w$ as the generalized path travel cost. We can transform Equation (5.34) into

$$\hat{f}_r^w = \exp[\frac{t_r^w - \mu_w}{-\hat{\theta}}], r \in R_w. \tag{5.35}$$

Solving for Equations (5.35) and (5.12), we can easily obtain

$$\hat{p}_r^w = \frac{\exp[\frac{t_r^w - \mu_w}{-\hat{\theta}}]}{\sum_{k \in R_w} \exp[\frac{t_k^w - \mu_w}{-\hat{\theta}}]} = \frac{\hat{f}_r^w}{\hat{q}_r^w}, r \in R_w, w \in W. \tag{5.36}$$

Equation (5.36) is just the SUE conditions for non-CAV users, which means that these drivers will choose their routes in accordance with the logit-based route choice probability. The same approach can be followed to show the equivalence of the route flow associated with the CAV users.

Therefore, the mixed SUE conditions are satisfied for the route flows associated with the CAV and non-CAV users where the Lagrangian multiplier ρ is the market equilibrium credit price. This completes the proof.

$\square$

Next, we will see that if such a price exists, it will be unique market clearing price.

Proposition 5.2.3. *(Uniqueness of market clearing price) Under the condition that the solution to the minimization programming problem exists and the credit market is clear, there exists a unique positive equilibrium.*

Proof. The existence of solution is guaranteed according to Propositions (5.2.1) and (5.2.2) as the objective function and the constraints are both convex for strictly increasing link performance functions. Moreover, we assumed that the credit scheme is effective such as that the obtained link flow pattern is not identical to the user equilibrium unique link flow pattern, i.e., $\sum_a \hat{\kappa}_a \hat{v}_a^0 + \bar{\kappa}_a \bar{v}_a^0 > K$ holds for the link flow pattern $(\hat{v}_a^0, \bar{v}_a^0)$ under no credit scheme intervention. At the market clearing price, we have $\sum_{a \in A} \hat{\kappa}_a \hat{v}_a + \bar{\kappa}_a \bar{v}_a = K$. Summing up Equations (5.28) and (5.30) over all $r \in R_w$, $w \in W$ and utilizing the conditions $\sum_{r \in R_w} \bar{f}_r^w = \bar{q}_w$ and $\sum_{r \in R_w} \hat{f}_r^w = \hat{q}_w$, $v_a = \sum_{w \in W} \sum_{r \in R^w} (\bar{f}_r^w + \hat{f}_r^w) \delta_{ar}^w$, $\hat{v}_a = \sum_{w \in W} \sum_{r \in R^w} (\hat{f}_r^w) \delta_{ar}^w$, $\bar{v}_a = \sum_{w \in W} \sum_{r \in R^w} (\bar{f}_r^w) \delta_{ar}^w$ we can obtain

$$\sum_{a \in A} (t_a(v_a) v_a + \rho(\hat{\kappa}_a \hat{v}_a + \bar{\kappa}_a \bar{v}_a)) + \hat{\theta} \sum_{w \in W} \sum_{r \in R_w} \hat{f}_r^w \ln \hat{f}_r^w + \bar{\theta} \sum_{w \in W} \sum_{r \in R_w} \bar{f}_r^w \ln \bar{f}_r^w$$
$$= \sum_{w \in W} (\lambda_w \bar{q}_w + \mu_w \hat{q}_w). \tag{5.37}$$

After substituting $\sum_{a \in A} (\hat{\kappa}_a \hat{v}_a + \bar{\kappa}_a \bar{v}_a) = K$, the equilibrium credit price under the scheme is then given by

$$\rho = \frac{\sum_{w \in W} (\lambda_w \bar{q}_w + \mu_w \hat{q}_w) - T}{\sum_{a \in A} (\hat{\kappa}_a \hat{v}_a + \bar{\kappa}_a \bar{v}_a)}, \tag{5.38}$$

where

$$T = \sum_{a \in A} t_a(v_a) v_a + \hat{\theta} \sum_{w \in W} \sum_{r \in R_w} \hat{f}_r^w \ln \hat{f}_r^w + \bar{\theta} \sum_{w \in W} \sum_{r \in R_w} \bar{f}_r^w \ln \bar{f}_r^w, \tag{5.39}$$

is the time equivalent of system-wide travel cost for the mixed stochastic traffic assignment problem. This completes the proof.

$\square$

Given our assumption on the monotonicity of the link performance function with the amount of traffic flow, we can see the convexity of the primal and thus the convexity of the dual problem. Therefore, from the dual problem perspective, we can also say that if such multipliers exists, then they must be unique minimal travel cost for each OD pair for each

group of travelers. We note that, the minimal path for each OD pair may not be unique, but the minimal cost for each OD pair, if exists, should be unique.

From Equation (5.37), we can see that the equilibrium market price (ρ), if exists, should be bounded. Therefore, for the case of a cyclic scheme, if we let $K \leq 0$, the equilibrium credit price still exists due to (5.37)) and can be given by

$$\rho = \lim_{K \to 0^-} \frac{\Theta - T}{K}, \tag{5.40}$$

where $\Theta = \sum_{w \in W} (\lambda_w \bar{q}_w + \mu_w \hat{q}_w)$ is the total OD travel disutility on the network.

This result is the extension of the findings on single-class deterministic settings (Xiao, Long, Li, Kou and Nie 2019, Yang and Wang 2011, Han and Cheng 2017) to multi-class stochastic environments. This is to make the proposed model self-sufficient and at the same time, it is also used as a basis for us to design our optimal compound credit-based schemes and solve it for optimal link-specific charge and reward rates on mixed-vehicle traffic networks in the following section.

Lemma 5.2.1. *Under perfect information scenario the market price of credits in a traffic network with fully CAVs is the same to the price of credits in deterministic traffic network.*

Proof. Under perfect information scenario for fully CAVs we will have $\hat{\theta} = \bar{\theta} \to 0$ and thus $T = \sum_{a \in A} t_a(v_a) v_a$. After setting $\lambda_w = \mu_w$ and $\hat{\kappa}_a = \bar{\kappa}_a = \kappa_a$ we can re-write Equation (5.38) as

$$\lim_{\hat{\theta} = \bar{\theta} \to 0} \rho = \frac{\sum_{w \in W} \mu_w q_w - \sum_{a \in A} t_a(v_a) v_a}{\sum_{a \in A} \kappa_a v_a}, \tag{5.41}$$

which is equivalent to the price of credits obtained for the deterministic case (Yang and Wang 2011). This completes the proof.

$\square$

5.2.3 Designing Optimal Compound Credit-based Scheme

Ideally, if all vehicles are under control and follow centralized travel decisions, the central authority could achieve a system-optimal state. Instead, we assume that the central authority uses a charging and subsidy scheme to improve the system efficiency by guiding it to move toward the so-called stochastic system-optimal state.

According to Hearn and Ramana (1998), when both system- and user-optimal equilibrium have unique solutions, one can find a decentralizing charging scheme located in the non-empty polyhedron defined in terms of given system-optimal solution. We can see that the first-order optimality conditions for the proposed model and thus the resulting mixed flow pattern is scale-invariant to the credit scheme. In other words, we can assume that the link flow and market equilibrium conditions remain unchanged if we uniformly scale up or down the total amount of credits and the link credit charges (Yang and Wang 2011). Thus, without loss of generality, we can set $\rho^* = 1$ in all of the following formulations.

In order to find a vector of the link-based charges in the context of SUE, which can give rise to the minimal total travel time on the network, we impose the first-best charges which can be obtained by solving an equivalent stochastic user-optimal problem with a modified objective function. In other words, the stochastic system-optimal (SSO) problem is equivalent to a standard SUE problem with link travel cost function defined as the summation of link travel time incurred by a traveler and the marginal travel time (additional travel time that the traveler imposes on all other travelers) in the link (Dafermos and Sparrow 1969), which is described as follows:

$$\tilde{t}_a(v_a) = t_a(v_a) + v_a \frac{\partial t_a(v_a)}{\partial v_a}, \tag{5.42}$$

where the first term on the right-hand side is the actual link travel time incurred by a traveler and the second term is the additional travel time that a traveler imposes on all other travelers in the link.

For a mixed (partially automated) traffic flow defined in Section (5.2.2), we set up the following charging system $(K, \hat{\kappa}, \bar{\kappa})$ that induces the stochastic system-optimal flow $v_a^* = \hat{v}_a + \bar{v}_a$

$$\sum_{a \in A}[(\sum_{w \in W}\sum_{r \in R_w}\hat{\theta}\ln \hat{f}^{*w}_r)\delta^w_{ar} + t_a(v^*_a) + \hat{\kappa}_a]\hat{v}^*_a = \sum_{w \in W}\hat{q}_w\mu_w, \tag{5.43}$$

$$\sum_{a \in A}(\sum_{w \in W}\sum_{r \in R_w}\hat{\theta}\ln \hat{f}^{*w}_r)\delta^w_{ar} + \sum_{a \in A}(t_a(v^*_a) + \hat{\kappa}_a)\delta^w_{ar} \geq \mu_w, r \in R_w, w \in W, \tag{5.44}$$

$$\sum_{a \in A}[(\sum_{w \in W}\sum_{r \in R_w}\bar{\theta}\ln \hat{f}^{*w}_r)\delta^w_{ar} + t_a(v^*_a) + \bar{\kappa}_a]\bar{v}^*_a = \sum_{w \in W}\bar{q}_w\lambda_w, \tag{5.45}$$

$$\sum_{a \in A}(\sum_{w \in W}\sum_{r \in R_w}\bar{\theta}\ln \bar{f}^{*w}_r)\delta^w_{ar} + \sum_{a \in A}(t_a(v^*_a) + \bar{\kappa}_a)\delta^w_{ar} \geq \lambda_w, r \in R_w, w \in W, \tag{5.46}$$

$$\sum_{a \in A}(\hat{\kappa}_a\hat{v}^*_a + \bar{\kappa}_a\bar{v}^*_a) = K, \tag{5.47}$$

where $\hat{\kappa}_a = \hat{v}_a\frac{\partial t_a(v_a)}{\partial v_a}$, $\bar{\kappa}_a = \bar{v}_a\frac{\partial t_a(v_a)}{\partial v_a}$ are set to be respective link charges equal to the monetary equivalent of marginal travel times for non-CAV drivers and CAV users on all other travelers on the link. This polyhedron is based on the linear system defined in Hearn and Yildirim (2002) and Yildirim and Hearn (2005), to characterize link-based toll pattern that can decentralize the given SO flow pattern. Generally, these set of conditions express that, for each user class in a system optimum state, the generalized travel cost including the credit charges and rewards should not be greater than the travel cost under the the UE condition for each path. Moreover, the overall network-wide transportation cost with credit charges and rewards on SO state should be equal to that of UE state. We also impose the credit feasibility condition trough Equation (5.47).

Proposition 5.2.4. *When the credit price is unique and positive, a system-optimal compound charging scheme* $(\hat{\kappa}, \bar{\kappa})$ *always exists with the credit price equal to 1.*

Proof. We let $T^* = \sum_{a \in A}t_a(v^*_a)v^*_a + \sum_{w \in W}\sum_{r \in R_w}(\hat{\theta}\hat{f}^{*w}_r\ln \hat{f}^{*w}_r + \bar{\theta}\bar{f}^{*w}_r\ln \bar{f}^{*w}_r)$, then if we multiply (5.43) -(5.46) with $\frac{T^*}{T^*+K}$, we can easily verify that

$$\sum_{a \in A}[(\sum_{w \in W}\sum_{r \in R_w}\hat{\theta}\ln \hat{f}^{*w}_r)\delta^w_{ar} + t_a(v^*_a) + \hat{\kappa}^*_a]\hat{v}^*_a = \sum_{w \in W}\mu^*_w\bar{q}_w, \tag{5.48}$$

$$\sum_{a \in A}(\sum_{w \in W}\sum_{r \in R_w}\hat{\theta}\ln \hat{f}^{*w}_r)\delta^w_{ar} + \sum_{a \in A}(t_a(v^*_a) + \hat{\kappa}^*_a)\delta^w_{ar} \geq \mu^*_w, r \in R_w, w \in W, \tag{5.49}$$

$$\sum_{a \in A}[(\sum_{w \in W}\sum_{r \in R_w}\bar{\theta}\ln \bar{f}^{*w}_r)\delta^w_{ar} + t_a(v^*_a) + \bar{\kappa}^*_a]\bar{v}^*_a = \sum_{w \in W}\lambda^*_w\bar{q}_w, \tag{5.50}$$

$$\sum_{a \in A}(\sum_{w \in W}\sum_{r \in R_w}\bar{\theta}\ln \bar{f}^{*w}_r)\delta^w_{ar} + \sum_{a \in A}(t_a(v^*_a) + \bar{\kappa}^*_a)\delta^w_{ar} \geq \lambda^*_w, r \in R_w, w \in W, \tag{5.51}$$

where

$$\hat{\kappa}_a^* = \frac{\hat{\kappa}_a T^* - K(t_a(v_a^*) + (\sum_{w \in W} \sum_{r \in R_w} \hat{\theta} \ln \hat{f}^*{}_r^w)\delta_{ar}^w)}{T^* + K}, \tag{5.52}$$

and

$$\bar{\kappa}_a^* = \frac{\bar{\kappa}_a T^* - K(t_a(v_a^*) + (\sum_{w \in W} \sum_{r \in R_w} \bar{\theta} \ln \bar{f}^*{}_r^w)\delta_{ar}^w)}{T^* + K}, \tag{5.53}$$

and

$$\mu_w^* = \frac{T^*}{T^* + K}\mu_w, \lambda_w^* = \frac{T^*}{T^* + K}\lambda_w. \tag{5.54}$$

We can also verify that

$$
\begin{aligned}
K &= \sum_{a \in A} \hat{\kappa}_a^* \hat{v}_a^* + \bar{\kappa}_a^* \bar{v}_a^* \\
&= \sum_{a \in A} \frac{\hat{\kappa}_a T^* - K(t_a(v_a^*) + (\sum_{w \in W} \sum_{r \in R_w} \hat{\theta} \ln \hat{f}^*{}_r^w)\delta_{ar}^w)}{T^* + K} \hat{v}_a^* \\
&+ \sum_{a \in A} \frac{\bar{\kappa}_a T^* - K(t_a(v_a^*) + (\sum_{w \in W} \sum_{r \in R_w} \bar{\theta} \ln \bar{f}^*{}_r^w)\delta_{ar}^w)}{T^* + K} \bar{v}_a^* \\
&= T^* \sum_{a \in A} \frac{\hat{\kappa}_a \hat{v}_a + \bar{\kappa}_a \bar{v}_a}{T^* + K} \\
&- \sum_{a \in A} \frac{\hat{v}_a^* K(t_a(v_a^*) + (\sum_{w \in W} \sum_{r \in R_w} \hat{\theta} \ln \hat{f}^*{}_r^w)\delta_{ar}^w)}{T^* + K} \\
&- \sum_{a \in A} \frac{\bar{v}_a^* K(t_a(v_a^*) + (\sum_{w \in W} \sum_{r \in R_w} \bar{\theta} \ln \bar{f}^*{}_r^w)\delta_{ar}^w)}{T^* + K} \\
&= \frac{T^* K - T^* K}{T^* + K} = 0. \tag{5.55}
\end{aligned}
$$

Given that for CAV users $\sum_{a \in A}(\sum_{w \in W} \sum_{r \in R_w} \bar{\theta} \ln \bar{f}^*{}_r^w)\bar{v}_a^* \delta_{ar}^w = \sum_{w \in W} \sum_{r \in R_w} \bar{\theta} \bar{f}^*{}_r^w \ln \bar{f}^*{}_r^w$, and that for non-CAV users $\sum_{a \in A}(\sum_{w \in W} \sum_{r \in R_w} \hat{\theta} \ln \hat{f}^*{}_r^w)\hat{v}_a^* \delta_{ar}^w = \sum_{w \in W} \sum_{r \in R_w} \hat{\theta} \hat{f}^*{}_r^w \ln \hat{f}^*{}_r^w$, after substitution and with a few simplifications the last two terms is reduced to KT^*. We

can see that K can be zero, however, link credit rates are not necessarily zero as we let the charges to be negative. This means that $(\hat{\kappa}_a^*, \bar{\kappa}_a^*)$ is a cyclic revenue-neutral compound charging scheme with $\rho = 1$ under which the total credit collected is zero. $\qquad\square$

5.2.4 Equivalent Path-free Model

The presented model in the previous section assumes that the road manager can enumerate all the paths that users may choose, which is not feasible for large-scale networks. In this section, we reformulate the proposed mathematical programming model to avoid path enumeration. To this end, we re-describe the network topology from a local point of view and represent a destination-based traffic flow model. The model presented here can also be used as a base model to model conditions such as link capacity constraints or demand elasticity.

Consider that the road network graph $G = (N, A)$, where $D \in N$ is a set of destinations nodes such that link flows directed toward specific destinations $d \in D$. Let us consider $V = \{V^1, V^2, ... V^{|D|}\}$ which each $V^d = \{v_{ij}^d | (i, j) \in A\}$ is the flow vector on link (i, j) destined to the destination $d = \{1, 2, ..., |D|\}$, with $\bar{v}_{ij}^d$ and $\hat{v}_{ij}^d$ to be, respectively, representative of flows for CAVs and non-CAVs on the corresponding link. Then the path-free mixed-vehicle stochastic user equilibrium traffic flow problem reads as the following convex minimization model

$$Min\ Z_d(\boldsymbol{\hat{v}}, \boldsymbol{\bar{v}}) = \sum_{(i,j)\in A} \int_0^{\hat{v}_{ij}+\bar{v}_{ij}} t_{ij}(x)\,dx$$

$$+\hat{\theta}(\sum_{d\in D}[\sum_{(i,j)\in A} \hat{v}_{ij}^d \ln \hat{v}_{ij}^d - \sum_{(i,j)\in A} \hat{v}_{ij}^d (ln \sum_{(i,k)\in A} \hat{v}_{ik}^d)])$$

$$+\bar{\theta}(\sum_{d\in D}[\sum_{(i,j)\in A} \bar{v}_{ij}^d \ln \bar{v}_{ij}^d - \sum_{(i,j)\in A} \bar{v}_{ij}^d (ln \sum_{(i,k)\in A} \bar{v}_{ik}^d)]) \tag{5.56}$$

subject to:

$$\sum_{j|(i,j)\in A} \hat{v}_{ij}^d - \sum_{j|(j,i)\in A} \hat{v}_{ji}^d = \hat{q}_i^d, i \in N, d \in D, \tag{5.57}$$

$$\sum_{j|(i,j)\in A} \bar{v}_{ij}^d - \sum_{j|(j,i)\in A} \bar{v}_{ji}^d = \bar{q}_i^d, i \in N, d \in D, \tag{5.58}$$

$$\hat{v}_{ij} = \sum_{d\in D} \hat{v}_{ij}^d, (i,j) \in A, \tag{5.59}$$

$$\bar{v}_{ij} = \sum_{d\in D} \bar{v}_{ij}^d, (i,j) \in A, \tag{5.60}$$

$$\sum_{ij\in A} \hat{\kappa}_{ij}\hat{v}_{ij} + \bar{\kappa}_{ij}\bar{v}_{ij} \le K, \tag{5.61}$$

$$\hat{v}_{ij} \ge 0, \bar{v}_{ij} \ge 0, (i,j) \in A. \tag{5.62}$$

Using this formulation, we then provide NCP equivalent formulations for the mixed stochastic user- and system-optimal traffic assignment problems.

NCP Equivalent UE Traffic Assignment Problem

The non-linear complementarity problem is an easy way to find stationary points for non-linear programs of a system of equations and inequalities (Karamardian 1969). Moreover, it helps us to avoid solving complex integration element of the objective function (5.56). We can establish an equivalent NCP formulation of the destination-based mixed stochastic user equilibrium traffic assignment problem (*MSUE-TAP*) model using following complementarity conditions:

MSUE-TAP

$$0 \leq \hat{v}_{ij}^d \perp \{t_{ij}(v_{ij}) + \hat{\theta}ln(\frac{\hat{v}_{ij}^d}{\sum_{(i,k)\in A} \hat{v}_{ik}^d}) + \pi_i^d - \pi_j^d\} \geq 0, (i,j) \in A, d \in D, \qquad (5.63)$$

$$0 \leq \pi_i^d \perp \{ \sum_{j|(i,j)\in A} \hat{v}_{ij}^d - \sum_{j|(j,i)\in A} \hat{v}_{ji}^d - \hat{q}_i^d\} \geq 0, i \in N, d \in D, \qquad (5.64)$$

$$0 \leq \hat{v}_{ij} \perp \{\hat{v}_{ij} - \sum_d \hat{v}_{ij}^d\} \geq 0, (i,j) \in A, \qquad (5.65)$$

$$0 \leq \bar{v}_{ij}^d \perp \{t_{ij}(v_{ij}) + \bar{\theta}ln(\frac{\bar{v}_{ij}^d}{\sum_{(i,k)\in A} \bar{v}_{ik}^d}) + \tau_i^d - \tau_j^d\} \geq 0, (i,j) \in A, d \in D, \qquad (5.66)$$

$$0 \leq \tau_i^d \perp \{ \sum_{j|(i,j)\in A} \bar{v}_{ij}^d - \sum_{j|(j,i)\in A} \bar{v}_{ji}^d - \bar{q}_i^d\} \geq 0, i \in N, d \in D, \qquad (5.67)$$

$$0 \leq \bar{v}_{ij} \perp \{\bar{v}_{ij} - \sum_d \bar{v}_{ij}^d\} \geq 0, (i,j) \in A, \qquad (5.68)$$

$$0 \leq v_{ij} \perp \{v_{ij} - \bar{v}_{ij} - \hat{v}_{ij}\} \geq 0, (i,j) \in A, \qquad (5.69)$$

where the component-wise notation "$\perp$" signifies that (at least) one of the inequalities must be satisfied as an equality. In this formulation, π_i^d and τ_i^d are minimum travel times from node i to destination node d under user-optimal solution for non-CAV and CAV users, respectively. The first constraint set (5.63) to (5.65) are complementarity conditions for non-CAV users. According to condition (5.63) only when the link (i,j) is on the shortest path from i to d we can have traffic flow on it. Condition set (5.64) for each link (i,j) are NCP equivalent of flow conservation constraints tied to their respective travel times. Condition set (5.64) for each (i,j) derive links flow of non-CAVs $\hat{v}_{ij}$ from their path flows $\hat{v}_{ij}^d$. Similar complementarity conditions are set through constraints (5.66) to (5.68) for CAV users. Finally, constraint set (5.69) ties the overall link flow v_{ij} to the sum of non-CAV and CAV link flows.

NCP Equivalent SO Traffic Assignment Problem

Given that the system-optimal objective function is to minimize the cost equivalent of total travel time which is given by

$$F(\hat{\boldsymbol{f}}, \bar{\boldsymbol{f}}) = \sum_{(i,j)\in A} t_{ij}(v_{ij})v_{ij} + \sum_{w\in W} \sum_{r\in R_w} (\hat{\theta}\hat{f}_r^w(\ln \hat{f}_r^w - 1) + \bar{\theta}\bar{f}_r^w(\ln \bar{f}_r^w - 1)). \qquad (5.70)$$

We build the following equivalent NCP model for the mixed stochastic system-optimal traffic assignment problem (*MSSO-TAP*) under a no charge scheme with modified condition sets (5.63) and (5.66) to include the marginal travel time cost for non-CAVs and CAVs, respectively.

MSSO-TAP

$$0 \leq \hat{v}_{ij} \perp \{t_{ij}(v_{ij}) + \hat{v}_{ij}\frac{\partial t_{ij}(v_{ij})}{\partial \hat{v}_{ij}} + \hat{\theta}ln(\frac{\hat{v}_{ij}^d}{\sum_{(i,k)\in A} \hat{v}_{ik}^d}) + \zeta_i^d - \zeta_j^d\} \geq 0, \qquad (5.71)$$

$$0 \leq \bar{v}_{ij} \perp \{t_{ij}(v_{ij}) + \bar{v}_{ij}\frac{\partial t_{ij}(v_{ij})}{\partial \bar{v}_{ij}} + \bar{\theta}ln(\frac{\bar{v}_{ij}^d}{\sum_{(i,k)\in A} \bar{v}_{ik}^d}) + \eta_i^d - \eta_j^d\} \geq 0, \qquad (5.72)$$

and (5.64), (5.65), (5.67), (5.68), (5.69),

where ζ_i^d and η_i^d respectively are minimum travel times from node i to destination node d for non-CAV and CAV users under a system-optimal solution. The rest of the complementarity conditions are similar to those in *MSUE-TAP*.

We note that, for the system optimal model the objective function (5.70) measures the net economic benefit defined as the travelers' benefit (the utility related to the aggregate demands) minus the total transportation cost. This is because the stochastic system optimal flow measures the sensitivity of route choices to total travel cost using the logit route choice model. According to Yang (1999), this measure has an economic interpretation related to consumer welfare. For example, if we ignore income effects, a change in price or any other characteristics of the travel environment results in a change in consumers' welfare.

5.2.5 A Set of Pareto-improving Compound Charging Schemes

When the objective of congestion charging scheme is to achieve a system-optimum equilibrium traffic flow pattern, in the context of SUE, stochastic system-optimum should be adopted as the objective of the charging scheme. In other words, an optimal charging pattern under a stochastic traffic flow pattern case has the same form as that which is optimal in the deterministic case, but evaluated at the SSO flow values instead of the Wardropian system-optimum flow values, even the ideal charging pattern may not be unique (Maher et al. 2005). Intuitively, one can see that the same condition would hold for finding a

94

charging scheme for the proposed mixed traffic model, as it is a generalization of SUE case. Given that the ideal charging pattern is not necessarily unique (Yang 1999), this enables us to choose a compound scheme $(K, \hat{\kappa}, \bar{\kappa})$ which satisfies certain exogenous constraints or objectives. In what follows we present two different charging schemes.

Permissive Scheme

To assure a revenue-neutral scheme we can assume $\sum_{(i,j) \in A} \hat{v}_{ij}^* \hat{\kappa}_{ij} + \bar{v}_{ij}^* \bar{\kappa}_{ij} = 0$. Moreover, as we discussed earlier, the transport authority may look for schemes under which the CAV users are subsidized more and charged less compared to non-CAV users. Such a differentiated charging scheme is suitable for the situations where an undifferentiated charging scheme might not be sufficient for achieving a system optimal state (Mehr and Horowitz 2019). This would also encourage CAV users and facilitate the acceptance of CAVs. In other words, we assume a charging scheme $\bar{\kappa}_{ij} \leq \hat{\kappa}_{ij}$, for all $(i, j) \in A$ such that the charge of traversing each link depends on whether vehicles are human–driven or CAVs.

Overall, we look for a (Pareto-improving) compound charging scheme and show that a stochastic socially optimum flow pattern is attainable by imposing discriminatory charges at each link to ensure that users' optimal private choices will also be optimal social choices in terms of the minimization of social cost. Our optimization model for finding link-based charging scheme is formulated as follows:

PO-1

$$Max \ \ Z(\hat{\boldsymbol{\kappa}}, \bar{\boldsymbol{\kappa}}) = \sum_{(i,j)\in A} \hat{\kappa}_{ij}\hat{v}_{ij}^* + \bar{\kappa}_{ij}\bar{v}_{ij}^*, \tag{5.73}$$

subject to:

$$\sum_{(i,j)\in A} [t_{ij}(v_{ij}^*) + \hat{\kappa}_{ij} + \hat{\theta}ln(\frac{\hat{v}_{ij}^*}{\sum_{(i,k)\in A} \hat{v}_{ik}^*})]\hat{v}_{ij}^* = \sum_{i\in N}\sum_{d\in D} \mu_i^d\hat{q}_i^d, \tag{5.74}$$

$$t_{ij}(v_{ij}^*) + \hat{\kappa}_{ij} + \hat{\theta}ln(\frac{\hat{v}_{ij}^{d,*}}{\sum_{(i,k)\in A} \hat{v}_{ik}^{d,*}}) + \mu_i^d - \mu_j^d \geq 0, (i,j) \in A, d \in D, \tag{5.75}$$

$$\sum_{(i,j)\in A} [t_{ij}(v_{ij}^*) + \bar{\kappa}_{ij} + \bar{\theta}ln(\frac{\bar{v}_{ij}^*}{\sum_{(i,k)\in A} \bar{v}_{ik}^*})]\bar{v}_{ij}^* = \sum_{i\in N}\sum_{d\in D} \lambda_i^d\bar{q}_i^d, \tag{5.76}$$

$$t_{ij}(v_{ij}^*) + \bar{\kappa}_{ij} + \bar{\theta}ln(\frac{\bar{v}_{ij}^{d,*}}{\sum_{(i,k)\in A} \bar{v}_{ik}^{d,*}}) + \lambda_i^d - \lambda_j^d \geq 0, (i,j) \in A, d \in D, \tag{5.77}$$

$$\bar{\kappa}_{ij} - \hat{\kappa}_{ij} \leq 0, (i,j) \in A, \tag{5.78}$$

$$\mu_i^d \leq \pi_i^d, i \in N, d \in D, \tag{5.79}$$

$$\lambda_i^d \leq \tau_i^d, i \in N, d \in D, \tag{5.80}$$

In this formulation the system-optimal flows $\hat{v}_{ij}^*$ and $\bar{v}_{ij}^*$ are given for each link. Variables μ_i^d and λ_i^d are the potentials of node i with respect to destination d for non-CAV and CAV users respectively, i.e., the length of path time (cost) from that node to destination node. With this definition, we can see that the potentials (cost) of the destination node should be zero ($\mu_d^d = 0, \lambda_d^d = 0$), and that μ_i^d and λ_i^d are the length of the minimum-cost path between i and destination d for the corresponding users. Constraints (5.75) and (5.77) ensure that the reduced cost of a link (i,j) for destination d not to be negative for non-CAVs and CAVs, respectively. In (5.79) and (5.80), π_i^d and τ_i^d are minimum travel time for non-CAVs and CAVs, respectively, and they ensure that the solution is Pareto-improving. These set of constraints along with constrains (5.74) - (5.77) guarantee the obtained charging scheme maintains the traffic flow under at the user-optimal level and the potential of each node is not greater that the obtained minimum travel times from that node to destination node.

The *PO-1* formulation can be replaced by the following equivalent problem, by introducing a slack variable $\xi \geq 0$.

PO-2

Min ξ $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (5.81)

subject to:

$$\sum_{(i,j)\in A} \hat{v}_{ij}^* \hat{\kappa}_{ij} + \bar{v}_{ij}^* \bar{\kappa}_{ij} = 0, \qquad\qquad\qquad (5.82)$$

$$\sum_{(i,j)\in A} [t_{ij}(v_{ij}^*) + \hat{\kappa}_{ij} + \hat{\theta}ln(\frac{\hat{v}_{ij}^*}{\sum_{(i,k)\in A} \hat{v}_{ik}^*})]\hat{v}_{ij}^* = \sum_{i\in N}\sum_{d\in D} \mu_i^d \hat{q}_i^d, \qquad (5.83)$$

$$t_{ij}(v_{ij}^*) + \hat{\kappa}_{ij} + \hat{\theta}ln(\frac{\hat{v}_{ij}^{d,*}}{\sum_{(i,k)\in A} \hat{v}_{ik}^{d,*}}) + \mu_i^d - \mu_j^d \geq 0, (i,j) \in A, d \in D, \qquad (5.84)$$

$$\sum_{(i,j)\in A} [t_{ij}(v_{ij}^*) + \bar{\kappa}_{ij} + \bar{\theta}ln(\frac{\bar{v}_{ij}^*}{\sum_{(i,k)\in A} \bar{v}_{ik}^*})]\bar{v}_{ij}^* = \sum_{i\in N}\sum_{d\in D} \lambda_i^d \bar{q}_i^d, \qquad (5.85)$$

$$t_{ij}(v_{ij}^*) + \bar{\kappa}_{ij} + \bar{\theta}ln(\frac{\bar{v}_{ij}^{d,*}}{\sum_{(i,k)\in A} \bar{v}_{ik}^{d,*}}) + \lambda_i^d - \lambda_j^d \geq 0, (i,j) \in A, d \in D, \qquad (5.86)$$

$$\bar{\kappa}_{ij} - \hat{\kappa}_{ij} \leq 0, (i,j) \in A, \qquad\qquad\qquad (5.87)$$

$$\mu_i^d \leq \xi\pi_i^d, i \in N, d \in D, \qquad\qquad\qquad (5.88)$$

$$\lambda_i^d \leq \xi\tau_i^d, i \in N, d \in D, \qquad\qquad\qquad (5.89)$$

$$\xi \geq 0. \qquad\qquad\qquad\qquad\qquad\qquad (5.90)$$

In this formulation, π_i^d and τ_i^d are minimum travel times from node i to destination node d for non-CAV and CAV users respectively under user-optimal solution, which are obtained from *MSUE-TAP* (5.63) to (5.69). The essential idea behind this formulation is based on the definition of "Pareto-improvement" for non-atomic user equilibrium problem. Next, in the spirit of Xiao, Long, Li, Kou and Nie (2019), we set the definition of Pareto-improvement as it applies to our model.

Definition 5.2.1. Under the compound credit-based scheme $(K, \bar{\kappa}, \hat{\kappa})$, if the system travel cost is $T(K, \bar{\kappa}, \hat{\kappa})$ and $\lambda_w(K, \bar{\kappa}, \hat{\kappa})$ and $\mu_w(K, \bar{\kappa}, \hat{\kappa})$ are the travel cost for CAVs and non-CAVs between OD pair w under the credit scheme $(K, \bar{\kappa}, \hat{\kappa})$. Then, the credit scheme $(K, \bar{\kappa}, \hat{\kappa})$ is Pareto-improving if the following conditions are satisfied:

- $T(K, \bar{\kappa}, \hat{\kappa}) \leq T(UE)$,

- $\lambda_w(K, \bar{\kappa}, \hat{\kappa}) \leq \lambda_w(UE)$,

- $\mu_w(K, \bar{\kappa}, \hat{\kappa}) \leq \mu_w(UE)$.

Given π_i^d and τ_i^d are minimum travel times, which are obtained from *MSUE-TAP*, in *PO-2* if at the optimal solution we have $\xi \leq 1$, then the credit scheme $(K, \bar{\kappa}, \hat{\kappa})$ is Pareto-improving. Otherwise, the Pareto-improving solution does not exist.

Lemma 5.2.2. *Under the credit-based scheme (5.73)-(5.80), the network contains no negative cycle.*

Proof. Here we only show that for non-CAV users no negative cycle exists, as for users with CAVs the proof would be similar. We assume P denotes the set of links of any directed cycle in the network. For any destination d, we have that

$$
\begin{aligned}
\sum_{(i,j)\in P} \tilde{t}_{ij}^d &= \sum_{(i,j)\in P} (t_{ij}(v_{ij}^*) + \hat{\kappa}_{ij} + \hat{\theta} ln(\frac{\hat{v}_{ij}^{d,*}}{\sum_{(i,k)\in A} \hat{v}_{ik}^{d,*}}) + \mu_i^d - \mu_j^d) \\
&= \sum_{(i,j)\in P} (t_{ij}(v_{ij}^*) + \hat{\kappa}_{ij} + \hat{\theta} ln(\frac{\hat{v}_{ij}^{d,*}}{\sum_{(i,k)\in A} \hat{v}_{ik}^{d,*}})) + \sum_{(i,j)\in P} (\mu_i^d - \mu_j^d) \\
&= \sum_{(i,j)\in P} (t_{ij}(v_{ij}^*) + \hat{\kappa}_{ij} + \hat{\theta} ln(\frac{\hat{v}_{ij}^{d,*}}{\sum_{(i,k)\in A} \hat{v}_{ik}^{d,*}})).
\end{aligned} \tag{5.91}
$$

From constraint (5.84), we have that

$$
\begin{aligned}
\sum_{(i,j)\in P} &(t_{ij}(v_{ij}^*) + \hat{\kappa}_{ij} + \hat{\theta} ln(\frac{\hat{v}_{ij}^{d,*}}{\sum_{(i,k)\in A} \hat{v}_{ik}^{d,*}})) \\
&= \sum_{(i,j)\in P} (t_{ij}(v_{ij}^*) + \hat{\kappa}_{ij} + \hat{\theta} ln(\frac{\hat{v}_{ij}^{d,*}}{\sum_{(i,k)\in A} \hat{v}_{ik}^{d,*}}) + \mu_i^d - \mu_j^d) \geq 0. \tag{5.92}
\end{aligned}
$$

Therefore, no negative cycle exist under the *PO-1* optimization problem. $\qquad\square$

Minimal Value (MV) Scheme

It may not be desirable if users are charged or subsidized with overly high rates. Essentially, this type of "choking" or high-pressure situation would raise social equity concerns. We

propose a charging scheme with the aim of avoiding excessively high values of charges or rewards. In this scheme we look for minimal charge and subsidy values by using the the minimal enclosing circle problem which is mostly used to plan the location of a shared facility with the aim of minimizing the distance between the facility and the farthest users (community) from it (Xu et al. 2003). If we think of the users on each link as communities in the network, we can minimize the excessively high charges/rewards by solving the following optimization problem.

PO-3

$$Min \ \ Z_d(\hat{\boldsymbol{\kappa}}, \bar{\boldsymbol{\kappa}}) = \sum_{(i,j) \in A} \hat{\kappa}_{ij}^2 \hat{v}_{ij}^* + \bar{\kappa}_{ij}^2 \bar{v}_{ij}^* \tag{5.93}$$

subject to:

$$(5.74), (5.75), (5.76), (5.77), (5.79), (5.80), and \ (5.82).$$

The only remaining issue is the existence of the feasible solution which can be readily put forward due to the fact that the credit values lie in a compact set and that the feasible set is convex because all the constraints are linear. Therefore, the set of feasible solutions is closed and bounded. The solution space is also non-empty provided that both *MSSO-TAP* and *MSUE-TAP* always have a solution. The only requirement for the objective function is to be integrable and positive, which holds true for both of the charging schemes. Hence problem *PO-1* and *PO-3* have a feasible solution.

The link-based compound charging scheme $(\hat{\kappa}, \bar{\kappa})$ under a mixed traffic flow can be found through following steps.

- **Step 1:** Solve the *MSUE-TAP*, and obtain OD travel dis-utilities π_i^d and τ_i^d;

- **Step 2:** Solve the *MSSO-TAP*, and obtain link flow $\hat{v}_{ij}^{*d}$, $\bar{v}_{ij}^{*d}$, and v_{ij}^*;

- **Step 3:** Solve *PO-2* or *PO-3*, and obtain the optimal $(K, \hat{\kappa}, \bar{\kappa})$. If the objective function for *PO-2* is less than 1 then the obtained scheme is Pareto-improving.
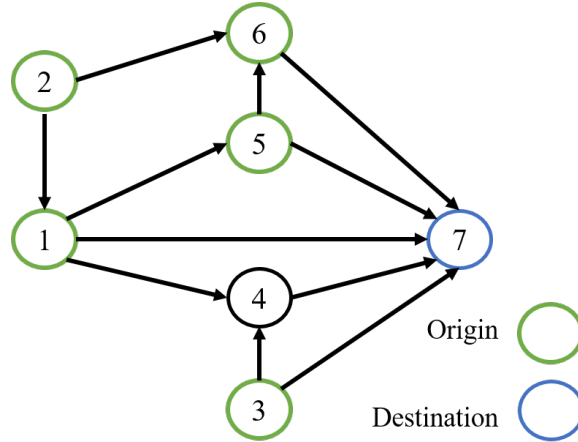
Figure 5.1 Toy Test Network with Multiple Origin and Single Destination

5.3 Numerical Results

5.3.1 Small-scale Networks

Our first test network is adopted from Liu, Wang and Meng (2014), to run numerical experiments and illustrate the performance of the proposed scheme. The test network has seven nodes, 11 links, five OD pairs ($1 \to 7$, $2 \to 7$, $3 \to 7$, $5 \to 7$, $6 \to 7$) with OD total CAV (non-CAV) vehicle per hour (v/h) travel demand, respectively, equal to 1000 (3000), 1000 (2000), 1000 (5000), 1000 (1000), 1000 (4000) (see Figure (5.1)). For each link, the link performance function takes the format proposed by the U.S. Bureau of Public Roads (BPR) (Spiess 1990) with $\alpha = 0.15$ and $\beta = 4$, for all $a \in A$. The value of free flow travel time and capacity of each link (v/h) are provided in Table (5.1). The marginal delay function for CAV and non-CAV flows are computed according to Equation (5.42). We assume that the values of dispersion parameters $\hat{\theta}$ and $\bar{\theta}$ in Equation (5.10) is 10 for non-CAV drivers and 2 for CAV drivers, implying a higher degree of uncertainty (5 times more) associated with non-CAV users (Yin and Yang 2003). The value of time in this test is taken as 1 cent/s (or 60\$/h).

For the sake of obtaining link-specific charge/subsidy rates, the path-free models of user- and system-optimal traffic assignment problems were solved. All experiments are run on a laptop computer with an Intel (R) Core(TM) 7 Oct 8550U 1.80 GHz CPU and a 15.8GB RAM. The optimal *MSUE-TAP* and *MSSO-TAP* link flows and link-specific charges are tabulated in Table (5.2) and (5.3). We found total travel time (*TTT*) cost to

Table 5.1 Parameters in Link Performance Functions

Link		t_a^0 (s)	C_a (v/h)
Tail	Head		
1	4	60	4000
1	7	150	4000
1	5	50	4000
2	1	60	2000
2	6	100	3000
3	4	70	3000
3	7	110	4000
4	7	60	4000
5	6	10	2000
5	7	70	4000
6	7	110	4000

be almost 51554 and 51052 for the computed *MSUE-TAP* and *MSSO-TAP*, respectively. This implies that with the proposed link-based scheme we can reduce the total system travel cost of this test example by one percent. For this small sized network, obtaining 1% improvement may not be very high, which can be due to the assumed setting and the value of parameters. Indeed, we might have some exceptional cases (e.g., small size problems) where the user equilibrium and system optimal flow patterns can lead to same total cost. One reason for this can be that we are not dealing with an overly congested network. However, it confirms that the proposed scheme have impact and improvement even on very small size networks.

Table 5.2 MSUE and MSSO Total Link Flows (v/h) and the Percentage of non-CAV and CAV Users ($\hat{\theta} = 20$ and $\bar{\theta} = 2$)

Link	MSUE-TAP			MSSO-TAP		
	Total (v/h)	non-CAVs (%)	CAVs (%)	Total (v/h)	non-CAVs (%)	CAVs (%)
(1,4)	3271.29	63.63	36.37	3021.39	62.72	37.28
(1,5)	2718.26	70.71	29.29	2542.07	69.29	30.71
(1,7)	863.73	98.43	1.57	1240.49	92.57	7.43
(2,1)	2853.27	64.95	35.05	2803.95	64.36	35.64
(2,6)	146.73	99.92	0.08	196.05	99.55	0.45
(3,4)	1422.86	99.33	0.67	1488.38	99.15	0.85
(3,7)	4577.15	78.36	21.64	4511.62	78.12	21.88
(4,7)	4694.14	74.46	25.54	4509.77	74.75	25.25
(5,6)	1.62	99.93	0.07	0.60	99.83	0.17
(5,7)	4716.64	61.91	38.09	4541.47	60.79	39.21
(6,7)	5148.35	80.58	19.42	5196.65	80.74	19.26
-	**TTT cost** = 51554			**TTT cost** = 51052		

Table 5.3 Three Pareto-improving Charging Schemes

Link	Non-permissive ($)		Permissive ($)		Minimal Value ($)	
	non-CAVs	CAVs	non-CAVs	CAVs	non-CAVs	CAVs
(1,4)	-0.89	-1.02	<-0.01	<-0.01	-0.27	-0.27
(1,5)	-0.68	-0.82	0.00	0.00	-0.27	-0.27
(1,7)	-2.26	-2.40	0.00	0.00	0.09	0.06
(2,1)	-1.56	-1.57	0.00	0.00	0.02	0.06
(2,6)	-1.27	-1.43	0.00	0.00	0.68	0.46
(3,4)	-0.97	-1.03	0.00	0.00	-0.27	-0.27
(3,7)	-2.22	-2.27	0.00	0.00	0.09	0.06
(4,7)	-1.24	-1.24	0.00	0.00	0.09	0.06
(5,6)	1.23	0.00	1.23	1.23	2.32	0.00
(5,7)	-1.45	-1.45	0.00	0.00	0.09	0.06
(6,7)	-2.61	-2.61	0.00	0.00	0.09	0.06

In this example, as the optimal charging is not unique, we present the results of three different charging policies. The first scheme considers a non-permissive charging (*PO-2* without condition (5.78)) policy, while the second scheme (PIP charges) considers permissive charging policy (i.e., $\bar{\kappa}_a \leq \hat{\kappa}_a$ for all $a \in A$) and the third scheme considers minimal value charging policy by solving *PO-3* problem.

Our second test problem is on a small-size multi-OD grid network, which is adopted from Yin and Yang (2003) with a few modifications to suit our setting. The test network has 12 nodes, 17 links, and five OD pairs ($1 \to 12, 2 \to 8, 2 \to 11, 5 \to 8, 5 \to 11$) with OD total CAV (non-CAV) demands, respectively, equal to 500(800), 150(400), 200(100), 200(200), 100(200) (see Figure (5.2)). For each link, the link performance function takes the BPR format with $\alpha = 0.15$ and $\beta = 4$, for all $a \in A$. The specific value of free flow travel time and capacity of each link are provided in Table (5.4) .

The values of dispersion parameters $\hat{\theta}$ and $\bar{\theta}$ in Equation (5.10) is taken as 20 for non-CAV drivers and 2 for CAV users. Similarly, the value of time is taken as 1 cent/s. Using the same computational resource, the optimal *MSUE-TAP* and *MSSO-TAP* link flows and
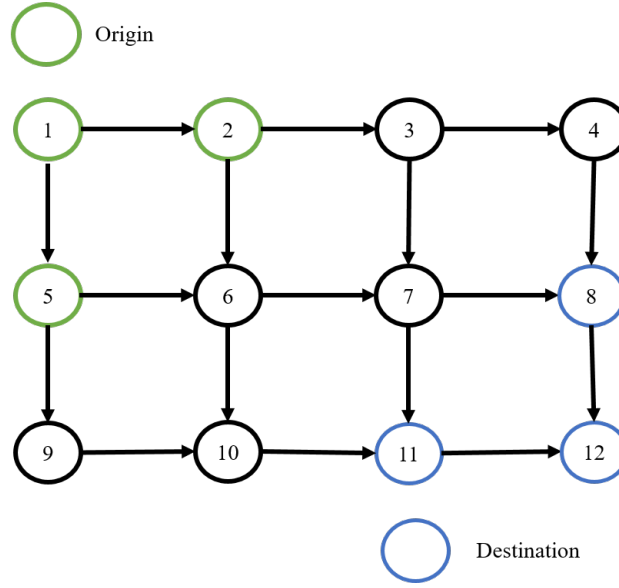
Figure 5.2 Multi Origin and Multi Destination Grid Network

link-specific charges are found and reported in Table (5.5) and (5.6).

We consider another set of scenarios to see how the uncertainty parameter will affect the link charge/subsidy values obtained from *PO-2* for both non-CAV and CAV users. To this end, we assume the OD demand values for both non-CAVs and CAVs to be equal for each OD pair. We then set the ratio of $y = \hat{\theta}/\bar{\theta}$ to be 1, 2, 3, 4, 8, and 10 while keeping $\bar{\theta} = 2$. A higher value of ratio implies a higher degree of non-CAVs uncertainty. The results are shown in Figure (5.3) and (5.4) for non-CAVs and CAVs, respectively.

It is interesting to see that in most case (for $y = 1, 2, 4$, and 10), as the value of uncertainty for non-CAVs increases, they will be charged more or subsidized less for the same link. However, when $y = 8$ in some links the proposed scheme turns to subside non-CAVs to maintain a Pareto-improving traffic equilibrium and reduce the system-wide transportation time. Interestingly, this is opposite to the general intuition that users with a higher value of uncertainty should be charged more or subsidize less to lead the network to a system-optimal state. On the other hand, we can see that the CAV users are charged more and subsidized less when $y = 8$, especially for link $(1, 2)$ and $(1, 5)$ while in other scenarios these users are mostly subsidized.

These findings drive us to check if we can find a charging scheme that can be more permissive towards CAV users while holding the Pareto-improving property. For this scenario, adding the side constrains $\bar{\kappa}_a \le \hat{\kappa}_a$ for all $a \in A$ and solving the same charging problem we

105

Table 5.4 Parameters of Link Performance Functions of the Grid Network

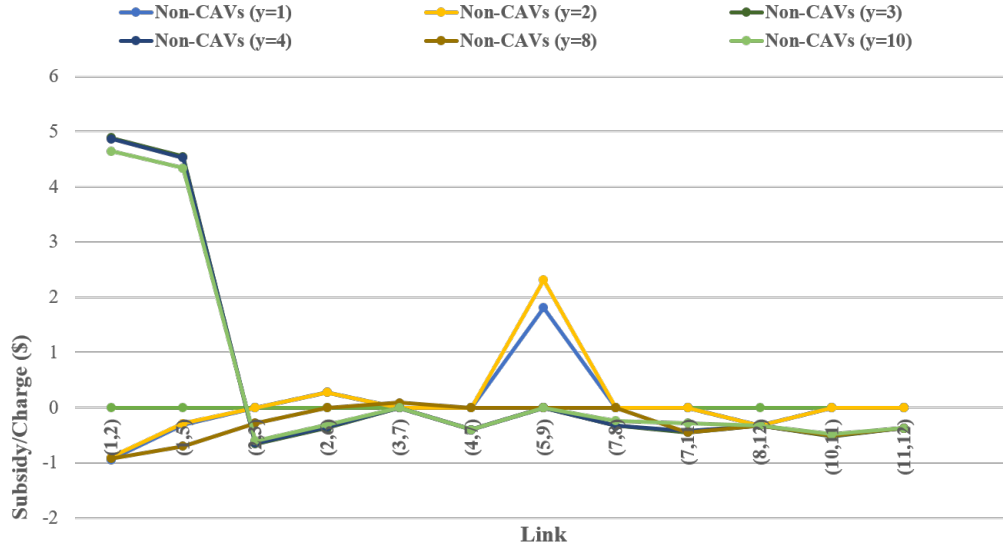| Link		t_a^0	C_a (v/h)	Link		t_a^0 (s)	C_a (v/h)
Tail	Head			Tail	Head		
1	2	20	1000	6	7	17	1000
1	5	18	1500	6	10	20	500
2	3	23	500	7	8	13	1000
2	6	19	500	7	11	26	500
3	4	17	500	8	12	19	1000
3	7	16	500	9	10	7	800
4	8	22	500	10	11	18	800
5	6	14	1000	11	12	17	800
5	9	24	800	-	-	-	-

Figure 5.3 Effect of Uncertainty Ratio (y) on Link Charge/Subsidy Values of non-CAVs

Table 5.5 MSUE and MSSO Link Flows (v/h) and TTT Costs $(\hat{\theta} = 20$ and $\bar{\theta} = 2)$

Link	MSUE-TAP (v/h)			MSSO-TAP (v/h)		
	Total	non-CAVs	CAVs	Total	non-CAVs	CAVs
(1, 2)	370.08	367.39	2.69	366.07	363.84	2.22
(1, 5)	929.92	432.61	497.31	933.93	436.15	497.78
(2, 3)	608.28	482.09	126.18	645.29	494.10	151.18
(2, 6)	611.80	385.29	226.50	570.77	369.73	201.03
(3, 4)	195.42	182.24	13.17	290.12	214.66	75.45
(3, 7)	412.85	299.85	113.00	355.17	279.44	75.73
(4, 8)	195.42	182.24	13.17	290.12	214.66	75.45
(5, 6)	1042.25	619.67	422.58	1007.82	618.18	389.63
(5, 9)	587.66	212.93	374.73	626.11	217.96	408.14
(6, 7)	1233.14	749.80	483.34	1199.81	744.17	455.64
(6, 10)	420.91	255.16	165.74	378.78	243.75	135.03
(7, 8)	1290.41	754.65	535.76	1174.96	726.23	448.72
(7, 11)	355.58	295.00	60.58	380.02	297.37	82.65
(8, 12)	535.83	336.89	198.95	515.08	340.90	174.17
(9, 10)	587.66	212.93	374.73	626.11	217.96	408.14
(10, 11)	1008.58	468.10	540.48	1004.89	461.72	543.17
(11, 12)	764.16	463.11	301.05	784.92	459.10	325.82
-	**TTT cost = 3836**			**TTT cost = 3780**		

Table 5.6 Two Pareto-improving Charging Schemes ($\hat{\theta} = 20$ and $\bar{\theta} = 2$)

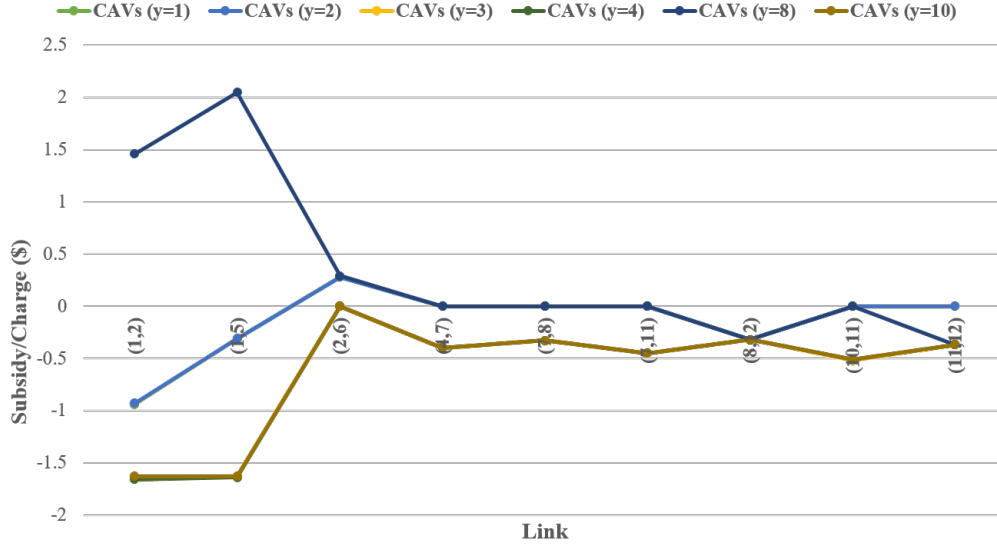| Link	Non-permissive ($)		Permissive ($)	
	non-CAVs	CAVs	non-CAVs	CAVs
(1,2)	3.24	-1.63	0.04	-0.68
(1,5)	2.94	-1.66	0.15	0.15
(2,3)	-0.45	0.00	0.00	0.00
(2,6)	-0.26	0.00	0.00	0.00
(3,4)	0.00	0.00	0.00	0.00
(3,7)	0.00	0.00	0.00	0.00
(4,8)	-0.34	-0.37	0.00	0.00
(5,6)	0.00	0.00	0.00	0.00
(5,9)	0.00	0.00	0.00	0.00
(6,7)	0.00	0.00	0.00	0.00
(6,10)	0.00	0.00	0.00	0.00
(7,8)	-0.21	-0.27	0.00	0.00
(7,11)	-0.26	-0.32	0.00	0.00
(8,12)	-0.32	-0.32	-0.32	-0.32
(9,10)	0.00	0.00	0.00	0.00
(10,11)	-0.37	-0.41	0.00	0.00
(11,12)	-0.32	-0.32	0.00	0.00

Figure 5.4 Effect of Uncertainty Ratio (y) on Link Charge/Subsidy Values of CAVs

find that $\hat{\kappa}_{(1,2)} = -0.03$, $\bar{\kappa}_{(1,2)} = -0.90$, $\hat{\kappa}_{(1,5)} = \bar{\kappa}_{(1,5)} = 0.19$, and $\hat{\kappa}_{(8,12)} = \bar{\kappa}_{(8,12)} = -0.32$ is another scheme that can sustain a Pareto-optimal equilibrium state.

Figure (5.5) plots the total travel time of the system under varied values of parameter y for *MSUE-TAP* and *MSSO-TAP* settings. It can be observed that as the uncertainty ratio increases the system travel time increases, which makes sense. However, total travel cost under user equilibrium doesn't change remarkably at some moderate uncertainty ratios ($y = 2$ to $y = 4$).

5.3.2 Medium-scale Network

Our next example is the classical Sioux Falls network which is an aggregation of a network used to model the city of Sioux Falls, South Dakota (LeBlanc 1975). It has 24 nodes, 76 links and 528 OD pairs. The link travel time functions are according to the standard BPR function with $\alpha = 0.15$ and $\beta = 4$, for all $a \in A$. The Sioux-Falls is a common benchmark case study network being used in the literature. The link performance data are provided in *https://github.com/bstabler/TransportationNetworks*. However, to adopt it for our setting, we assumed that travel demand of non-CAV users is 1.5 times the original demand data and travel demand from CAV users is 0.5 times the original travel demand. Also, the values of dispersion parameters $\hat{\theta}$ and $\bar{\theta}$ is taken as 20 for non-CAV drivers and
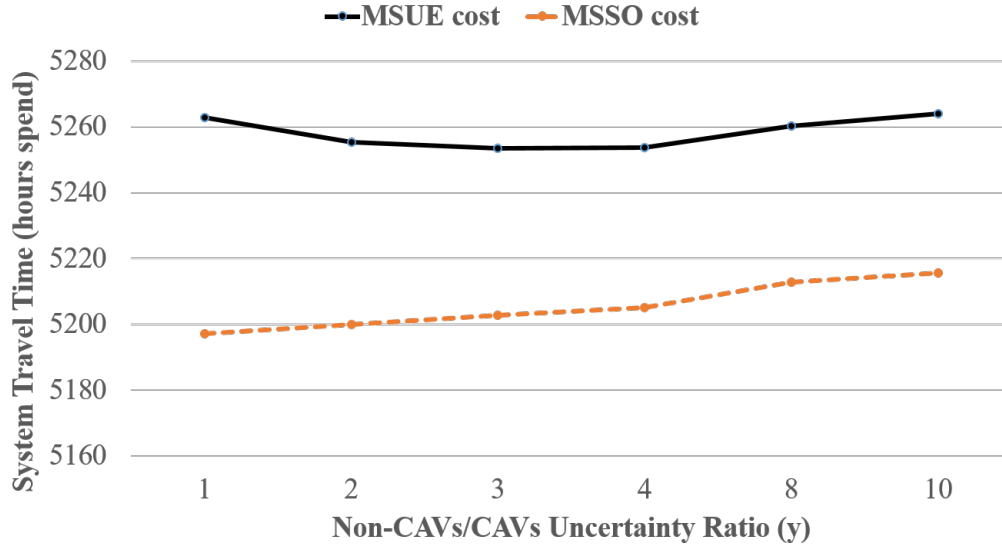
Figure 5.5 Effect of Uncertainty Ratio (y) on TTT

2 for CAV users. By solving the *MSUE-TAP* and *MSSO-TAP*, we have TTT cost to be 217792.37 and 206856.55 for the user-and system-optimal traffic flow patterns. This implies that a traffic flow under a Pareto-improving charging scheme can reduce the TTT cost by 5%. The performance of the proposed schemes (IP, PIP, MV) is compared in Table (5.7). One can observe that under the MV-charging scheme charge rates are much smaller than those under the other two, with CAV users are being subsidized in all links. While all the proposed charging schemes can reduce the total system travel cost for this problem, there might be other charging schemes as well, depending on the authority's objectives. The final choice would be a managerial decision.

We now turn to the analysis of the computational effort required to run the proposed scheme in this study. The credit finding algorithm and all the numerical experiments are solved in GAMS which is a widely used optimization software to model and solve many transport planning problems. All experiments are run on a laptop computer with an Intel (R) Core(TM) 7 Oct 8550U 1.80 GHz CPU and a 15.8GB RAM. Generally, the computational effort grows as the network size or number of OD pairs increases for the same network. In Sioux Falls network, which is the largest test problem we solved, it takes less than 300 s to solve the credit finding problem.
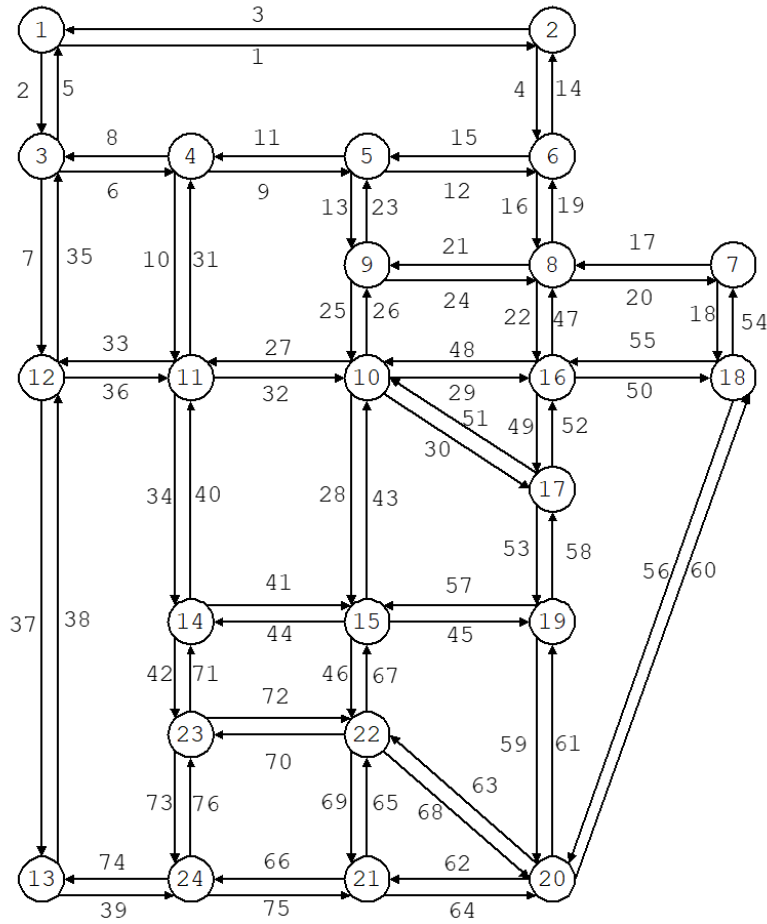
Figure 5.6 The Sioux Falls Network

Table 5.7 Performance of the Proposed Schemes for Sioux Falls Network ($\hat{\theta} = 16$ and $\bar{\theta} = 2$)

Scheme	Fleet Type	Max. ($)	Min. ($)	SD. ($)	Links Charged	Links Subsidized
Non-permissive	non-CAVs	12.12	-6.12	3.28	24	45
	CAVs	0	-6.29	1.40	0	71
Permissive	non-CAVs	31.07	-6.29	4.44	19	48
	CAVs	6.55	-6.29	2.23	17	50
Minimal Value	non-CAVs	1.57	-0.23	0.46	16	60
	CAVs	-0.23	-0.23	0.00	0	76

5.4 Conclusion

In this chapter, we present a compound credit-based scheme for managing traffic flow composed of CAVs and non-CAVs, considering characteristics such as AV users cooperative versus selfish routing principles and the uncertainty due to operational setting, human associated factors and their interaction due to using shared infrastructures such as network links. We develop logit-based mixed-vehicle stochastic user- and system-optimal traffic flow problems as mathematical models with non-linear complementarity inequalities. One of the contributions of this work is to build a theoretical ground to obtain insights pertaining to the assumed setting. Under the proposed scheme, we examine the existence and uniqueness of the equilibrium price of the credit and look for a set of link-specific charges/subsidies that can drive the network-wide traffic flow to an optimum state.

The proposed link-based congestion charging framework is an initial step for managing mixed-vehicle traffic flow. To implement the scheme in a real-world setting, many interesting issues are still to be explored. We performed computational experiments on the well-known datasets to analyze the performance of the proposed models. We tested all models under various different parameter settings. The proposed scheme relies on solving link-based convex optimization programming models, obviating computationally intractable path enumeration. This makes the proposed model suitable for examining the theoretical characteristics of large-scale as well realistic transport networks. The computational results indicate that the proposed scheme can be an effective tool for managing travel demand and routing decisions in mixed-vehicle traffic settings.

Chapter 6

Conclusions and Future Work

This chapter provides a summary and the main contributions of the thesis as well as future research directions.

6.1 Summary of Research Contributions

The main goal of this research was to design and examine the potential of mobility permit-based and credit-based traffic congestion management systems. Specifically, the research focused on developing new models and solution algorithms for permit allocation under single-bottleneck roadways as well as networks with multiple bottlenecks settings under advanced connectivity technologies.

In Chapter 2, we examined the state-of-the-art in permit-based and credit-based mobility management systems. A detailed and systematic literature review was conducted to identify previously proposed models and solution approaches. This has provided new insights into permit-based and credit-based mobility systems, their main advantages and limitations, laying foundation for future research in this field.

Next, In Chapter 3 we developed an integrated framework to identify and integrate the interests of all major stakeholders in a permit-based mobility market, such as mobility users, mobility service providers, system regulator and the general public. The main objective was to provide a road-map that covers the key elements of a futuristic user-centric traffic management system and then incorporate their requirements into a mobility regulatory scheme. This research has contributed to the literature by proposing the conceptual

framework of an integrated user-centric mobility permit-based traffic management system with particular attention on developing new pricing and allocation protocols under fairness and efficiency requirements.

In Chapter 4, we focused on the design and analysis of pricing and allocation of mobility permits for roadways with one bottleneck. We dealt with observing operational objectives, particularly, balancing efficiency and fairness in mobility permit allocation. We then explored the theoretical properties of the proposed scheme and showed that the proposed scheme can achieve an optimal traffic pattern; though, it is computationally intensive to solve large size problem instances. Next, to tackle the computational complexity of the proposed scheme, we proposed a heuristic permit allocation algorithm that sustains a Pareto-optimal traffic pattern with less computational effort. Next, we designed an iterative auction mechanism for pricing the mobility permit under two different pricing methods. To analyze the performance of the proposed schemes, we performed comprehensive numerical experiments under different parameter settings. We showed that a hybrid mechanism with a minimal over-demanded set pricing and heuristic allocation method can be a good candidate for being the mobility scheme component of the proposed integrated user-centric traffic management system. The presented MP-based traffic management scheme is Pareto-improving which increases its potentials to be economically and socially acceptable. Our main contribution was to incorporate revenue and fairness aspects into designing a single-bottleneck mobility management scheme while relaxing the assumption of availability of full information in determining the optimal price and allocation of permits to manage travel demand.

In Chapter 5, we developed models and algorithms for pricing and allocation of permits in a user-centric permit-based mobility system on single-bottleneck roadways. In this regard, considering the underlying issues, we presented new models and algorithms for permit pricing and allocation problems with an explicit consideration of equity and efficiency requirements. Moreover, we provided some properties related to the existence of equilibrium and optimality of solution, and investigate the effect of various assumptions. We then presented a compound credit-based scheme for managing traffic flow composed of human-driven and CAVs with ATIS, considering characteristics of such networks such as driver behavior related to human-driven vehicles and the uncertainty associated with the interaction between human-driven and CAVs. We developed logit-based mixed-vehicle stochastic user- and system-optimal traffic flow problems as mathematical models with non-linear complementarity inequalities. One of the contributions of this chapter was to

114

build a theoretical ground to obtain theoretical insights pertaining to the assumed setting. Under the proposed scheme, we examined the existence and uniqueness of the equilibrium price of the credit and look for a set of link-specific charges/subsidies that can drive the network-wide traffic flow to an optimum state. Another contribution of this chapter from a modeling perspective was that we developed novel models for mixed-fleet stochastic user- and system-optimal traffic assignments as non-linear complementarity problems (NCPs) and use them to find (Pareto-improving) link-specific charges. This means that the proposed scheme relies on solving link-based convex optimization programming models, obviating computationally intractable path enumeration. This makes the proposed model suitable for examining the theoretical characteristics of large-scale as well realistic networks. The computational results indicate that the proposed scheme can be an effective tool for managing travel demand and routing decisions in mixed-vehicle traffic settings. Our main contribution lies in improving traffic in mixed–vehicle networks, i.e., removing user equilibria inefficiency, via a compound credit-based charge and reward scheme. In addition, we studied the problem of assigning different charge and reward rates to network links for inducing a system-optimal traffic state under different schemes. Our findings supported that a mixed-vehicle traffic network can be led to a system-optimum state using a compound charge and reward travel demand management scheme.

Findings from this research can be useful to support the decisions of mobility service providers, city planners, or transport authorities who are looking to implement initiatives for managing travel demand and promoting certain travel behavior within conventional transportation networks. The managerial insights are not only for policy makers, but also for private sector transport service providers.

6.2 Recommendations for Future Work

One of the major challenges arising in designing a mobility scheme is how to balance the conflicting interests of competing stakeholders and participants of the system. For instance, fairness and economical viability have been characterised as key issues with significant effects in common resource allocation settings. The simplicity and convenience for individual traveller, and computational efficiency of the solution method are other implementation-related factors that can highly affect the realization of the roadway-use management schemes.

In practice, network operational capacity are highly governed by the network characteristics such as moving bottlenecks and route overlapping. A more complicated situation would be considering all these settings together in the design of traffic control approaches and algorithms for mixed traffic conditions. To account for the route overlapping and other user- and vehicle-oriented uncertainties, we simply use a variability parameter in the objective function; a different approach can be using a modified-logit route-choice model such as cross nested-logit with a coefficient component accounting for the covariance between different paths. However, this would entail a path enumeration process that scales up the computational complexity of the problem for a large-size network. No study considered demand uncertainty and supply variability simultaneously within a scheme design. Some previous work has focused on one of these uncertainties (Han and Cheng 2017), but not in combination. Networks with variable bottlenecks where the route (link) capacities and travel demand vary with certain probabilities or over time can be considered in future research. The findings in this study can be extended to some general directions by considering characteristics of dynamic traffic networks such user elasticity and heterogeneity, incorporating link capacity, and simultaneous demand and capacity fluctuations.

We proposed permissive, non-permissive, and minimal value schemes as different possible charging policies that can be implemented under specific settings. Future extension can be considering specific features of CAVs such as fuel-saving and/or high travel efficiency (higher average speed) in designing more specific charging schemes. Schemes with different charging and pricing schemes such as discount and bundle options also need to be investigated. Another problem that is worth investigating is the efficiency loss of a scenario where CAVs are not managed as a fleet. In this scenario, the credit scheme might be applied (to CAVs only) as a decentralized way to achieve a system optimum (possibly Pareto-improving). Moreover, our framework might be applied with adjustments to other settings such as CAV-restricted scenario where the CAV users would follow centralized travel recommendations to meet certain system-wide objectives.

References

Akamatsu, T., Sato, S. and Nguyen, L. (2006), 'Tradable time-of-day bottleneck permits for morning commuters', *JSCE Journal of Infrastructure Planning and Management* **62**(4), 605–620.

Akamatsu, T. and Wada, K. (2017), 'Tradable network permits: A new scheme for the most efficient use of network capacity', *Transportation Research Part C: Emerging Technologies* **79**, 178–195.

Azevedo, E. M. and Budish, E. (2017), Strategy-proofness in the large, Technical report, National Bureau of Economic Research.

Babones, S. (2018), 'Robo-taxis are the future of transportation—and china's didi is racing to get there first'.

Bao, Y., Gao, Z., Xu, M. and Yang, H. (2014), 'Tradable credit scheme for mobility management considering travelers' loss aversion', *Transportation Research Part E: Logistics and Transportation Review* **68**, 138–154.

Bao, Y., Gao, Z., Yang, H., Xu, M. and Wang, G. (2017), 'Private financing and mobility management of road network with tradable credits', *Transportation Research Part A: Policy and Practice* **97**, 158–176.

Bao, Y., Gaoa, Z. and Xu, M. (2016), 'Traffic assignment under tradable credit scheme: An investigation considering travelers' framing and labeling of credits', *IEEE Intelligent Transportation Systems Magazine* **8**(2), 74–89.

Bertoncello, M. and Wee, D. (2015), 'Ten ways autonomous driving could redefine the automotive world', *McKinsey & Company* **6**, (Accessed 12/10/2017).

Bertsimas, D., Farias, V. F. and Trichakis, N. (2012), 'On the efficiency-fairness trade-off', *Management Science* **58**(12), 2234–2250.

Cervero, R. (2017), 'Mobility niches: jitneys to robo-taxis', *Journal of the american planning association* **83**(4), 404–412.

Chen, L. and Yang, H. (2012), 'Managing congestion and emissions in road networks with tolls and rebates', *Transportation Research Part B: Methodological* **46**(8), 933–948.

Chen, S., Wang, H. and Meng, Q. (2019), 'Designing autonomous vehicle incentive program with uncertain vehicle purchase price', *Transportation Research Part C: Emerging Technologies* **103**, 226–245.

Chen, Z., Lin, X., Yin, Y. and Li, M. (2020), 'Path controlling of automated vehicles for system optimum on transportation networks with heterogeneous traffic stream', *Transportation Research Part C: Emerging Technologies* **110**, 312–329.

Commission, C. E. et al. (2015), 'We can't get there from here: Why pricing traffic congestion is critical to beating it'.

Cramton, P., Shoham, Y. and Steinberg, R. (2006), *Combinatorial Auctions*, MIT Press. **URL:** *https://books.google.ca/books?id=N_IYAQAAIAAJ*

Crocker, T. D. (1966), 'The structuring of atmospheric pollution control systems', *The economics of air pollution* **61**, 81–84.

Dafermos, S. C. and Sparrow, F. T. (1969), 'The traffic assignment problem for a general network', *Journal of Research of the National Bureau of Standards B* **73**(2), 91–118.

Dales, J. H. (1968), 'Land, water, and ownership', *The Canadian Journal of Economics/Revue canadienne d'Economique* **1**(4), 791–804.

Davis, L. W. (2008), 'The effect of driving restrictions on air quality in mexico city', *Journal of Political Economy* **116**(1), 38–81.

Demange, G., Gale, D. and Sotomayor, M. (1986), 'Multi-item auctions', *Journal of Political Economy* **94**(4), 863–872.

Doan, K., Ukkusuri, S. and Han, L. (2011), 'On the existence of pricing strategies in the discrete time heterogeneous single bottleneck model', *Transportation Research Part B: Methodological* **45**(9), 1483–1500.

Dogterom, N. (2017), Tradable Driving Credits: car users' responses towards an innovative pricing measure in the Netherlands and China, PhD thesis, Utrecht University.

Downs, A. (1962), 'The law of peak-hour expressway congestion', *Traffic Quarterly* **16**(3).

Fan, W. and Jiang, X. (2013), 'Tradable mobility permits in roadway capacity allocation: Review and appraisal', *Transport Policy* **30**, 132–142.

Force, U. T. T. (2012), 'The high cost of congestion in canadian cities', *Council of Ministers Transportation and Highway Safety* .

Foxon, T. J. (2011), 'A coevolutionary framework for analysing a transition to a sustainable low carbon economy', *Ecological Economics* **70**(12), 2258–2267.

Gao, G. and Hu, J. (2015), 'Optimal tradable credits scheme and congestion pricing with the efficiency analysis to congestion', *Discrete Dynamics in Nature and Society* **2015**.

Gao, G. and Sun, H. (2014), 'Internalizing congestion and emissions externality on road networks with tradable credits', *Procedia-Social and Behavioral Sciences* **138**, 214–222.

Gao, G., Sun, H., Wu, J., Liu, X. and Chen, W. (2018), 'Park-and-ride service design under a price-based tradable credits scheme in a linear monocentric city', *Transport Policy* **68**, 1–12.

Gao, G., Sun, H., Wu, J. and Zhao, H. (2016), 'Tradable credits scheme and transit investment optimization for a two-mode traffic network', *Journal of Advanced Transportation* **50**(8), 1616–1629.

Garey, M. R. and Johnson, D. S. (1990), *Computers and Intractability; A Guide to the Theory of NP-Completeness*, W. H. Freeman & Co., New York, NY, USA.

Gärling, T., Eek, D., Loukopoulos, P., Fujii, S., Johansson-Stenman, O., Kitamura, R., Pendyala, R. and Vilhelmson, B. (2002), 'A conceptual analysis of the impact of travel demand management on private car use', *Transport Policy* **9**(1), 59–70.

Godard, O. (2001), *Domestic transferable permits for environmental management: design and implementation*, Organization for Economic.

Goddard, H. C. (1997), 'Using tradeable permits to achieve sustainability in the world's large cities: policy design issues and efficiency conditions for controlling vehicle emissions, congestion and urban decentralization with an application to mexico city', *Environmental and Resource Economics* **10**(1), 63–99.

Grant-Muller, S. and Xu, M. (2014), 'The role of tradable credit schemes in road traffic congestion management', *Transport Reviews* **34**(2), 128–149.

Gu, Y., Deakin, E. and Long, Y. (2017), 'The effects of driving restrictions on travel behavior evidence from beijing', *Journal of Urban Economics* **102**, 106–122.

Gulipalli, P. K. and Kockelman, K. M. (2008), 'Credit-based congestion pricing: A dallas-fort worth application', *Transport Policy* **15**(1), 23–32.

Hall, P. (1935), 'On representatives of subsets', *Journal of the London Mathematical Society* **1**(1), 26–30.

Han, D., Yang, H. and Wang, X. (2010), 'Efficiency of the plate-number-based traffic rationing in general networks', *Transportation Research Part E: Logistics and Transportation Review* **46**(6), 1095–1110.

Han, F. and Cheng, L. (2015*a*), Stochastic user equilibrium traffic assignment problem with tradable credit scheme: Model and algorithm, *in* 'CICTP 2015', pp. 3897–3909.

Han, F. and Cheng, L. (2015*b*), Tradable credit scheme for network congestion management: Considering travelers' perception errors, *in* 'CICTP 2015', pp. 4003–4014.

Han, F. and Cheng, L. (2017), 'Stochastic user equilibrium model with a tradable credit scheme and application in maximizing network reserve capacity', *Engineering Optimization* **49**(4), 549–564.

Harker, P. T. (1988), 'Multiple equilibrium behaviors on networks', *Transportation science* **22**(1), 39–46.

Haurie, A. and Marcotte, P. (1985), 'On the relationship between nash—cournot and wardrop equilibria', *Networks* **15**(3), 295–308.

He, F., Yin, Y., Shirmohammadi, N. and Nie, Y. M. (2013), 'Tradable credit schemes on networks with mixed equilibrium behaviors', *Transportation Research Part B: Methodological* **57**, 47–65.

Hearn, D. W. and Ramana, M. V. (1998), Solving congestion toll pricing models, *in* 'Equilibrium and advanced transportation modelling', Springer, pp. 109–124.

Hearn, D. W. and Yildirim, M. B. (2002), A toll pricing framework for traffic assignment problems with elastic demand, *in* 'Transportation and network analysis: Current trends', Springer, pp. 135–145.

Hooker, J. N. and Williams, H. P. (2012), 'Combining equity and utilitarianism in a mathematical programming model', *Management Science* **58**(9), 1682–1693.

Hu, X., Chiu, Y.-C. and Zhu, L. (2015), 'Behavior insights for an incentive-based active demand management platform', *International Journal of Transportation Science and Technology* **4**(2), 119–133.

Johnston, R. A., Lund, J. R. and Craig, P. P. (1995), 'Capacity-allocation methods for reducing urban traffic congestion', *Journal of Transportation Engineering* **121**(1), 27–39.

Karamardian, S. (1969), 'The nonlinear complementarity problem with applications, part 1', *Journal of Optimization Theory and Applications* **4**(2), 87–98.

Karsu, Ö. and Morton, A. (2015), 'Inequity averse optimization in operational research', *European journal of operational research* **245**(2), 343–359.

Khattak, A. J. and Yim, Y. (2004), 'Traveler response to innovative personalized demand-responsive transit in the san francisco bay area', *Journal of urban planning and development* **130**(1), 42–55.

Klein, I. and Ben-Elia, E. (2016), 'Emergence of cooperation in congested road networks using ict and future and emerging technologies: A game-based review', *Transportation Research Part C: Emerging Technologies* **72**, 10–28.

Kockelman, K. M. and Kalmanje, S. (2005), 'Credit-based congestion pricing: a policy proposal and the public's response', *Transportation Research Part A: Policy and Practice* **39**(7), 671–690.

Koolstra, K. (1999), Potential benefits of a freeway slot-reservation system: Queuing costs versus scheduling costs.

Kristoffersson, I., Engelson, L. and Börjesson, M. (2017), 'Efficiency vs equity: Conflicting objectives of congestion charges', *Transport Policy* **60**, 99–107.

Lahlou, S. and Wynter, L. (2017), 'A nash equilibrium formulation of a tradable credits scheme for incentivizing transport choices: From next-generation public transport mode choice to hot lanes', *Transportation Research Part B: Methodological* **101**, 185–212.

LeBlanc, L. J. (1975), 'An algorithm for the discrete network design problem', *Transportation Science* **9**(3), 183–199.

Lessan, J. and Karabatı, S. (2018), 'A preference-based, multi-unit auction for pricing and capacity allocation', *Computers & Operations Research* **91**, 237–246.

Levin, M. W. and Boyles, S. D. (2015), 'Effects of autonomous vehicle ownership on trip, mode, and route choice', *Transportation Research Record* **2493**(1), 29–38.

Levin, M. W. and Boyles, S. D. (2016), 'A multiclass cell transmission model for shared human and autonomous vehicle roads', *Transportation Research Part C: Emerging Technologies* **62**, 103–116.

Levinson, D. (2010), 'Equity effects of road pricing: A review', *Transport Reviews* **30**(1), 33–57.

Li, Q. and Gao, Z. (2014), 'Managing rush hour congestion with lane reversal and tradable credits', *Mathematical Problems in Engineering* **2014**.

Li, R., Liu, X. and Nie, Y. M. (2018), 'Managing partially automated network traffic flow: Efficiency vs. stability', *Transportation Research Part B: Methodological* **114**, 300–324.

Li, Y., Ukkusuri, S. V. and Fan, J. (2018), 'Managing congestion and emissions in transportation networks with dynamic carbon credit charge scheme', *Computers & Operations Research* **99**, 90–108.

Litman, T. (2002), 'Evaluating transportation equity', *World Transport Policy & Practice* **8**(2), 50–65.

Liu, B. and Huang, C. (2014), 'Alternative tradable credit scheme on general transportation network', *Wuhan University Journal of Natural Sciences* **19**(3), 193–198.

Liu, W., Yang, H. and Yin, Y. (2015), 'Efficiency of a highway use reservation system for morning commute', *Transportation Research Part C: Emerging Technologies* **56**, 293–308.

Liu, W., Yang, H., Yin, Y. and Zhang, F. (2014), 'A novel permit scheme for managing parking competition and bottleneck congestion', *Transportation Research Part C: Emerging Technologies* **44**, 265–281.

Liu, Y. and Nie, Y. M. (2017), 'A credit-based congestion management scheme in general two-mode networks with multiclass users', *Networks and Spatial Economics* **17**(3), 681–711.

Liu, Z., Wang, S. and Meng, Q. (2014), 'Toll pricing framework under logit-based stochastic user equilibrium constraints', *Journal of Advanced Transportation* **48**(8), 1121–1137.

Lo, H. K. and Szeto, W. (2002), 'A methodology for sustainable traveler information services', *Transportation Research Part B: Methodological* **36**(2), 113–130.

Luo, Q., Saigal, R., Chen, Z. and Yin, Y. (2019), 'Accelerating the adoption of automated vehicles by subsidies: A dynamic games approach', *Transportation Research Part B: Methodological* **129**, 226–243.

Maher, M. and Hughes, P. (1995), Estimation of the potential benefits from an att system using a multiple user class stochastic user equilibrium assignment model, *in* 'Applications of advanced technologies in transportation engineering', ASCE, pp. 700–704.

Maher, M., Stewart, K. and Rosa, A. (2005), 'Stochastic social optimum traffic assignment', *Transportation Research Part B: Methodological* **39**(8), 753–767.

Mayeres, I. and Proost, S. (2002), 'Reforming transport pricing: an economist's perspective on equity, efficiency and acceptability'.

Mehr, N. and Horowitz, R. (2019), 'Pricing traffic networks with mixed vehicle autonomy', *CoRR* **abs/1904.01226**.
URL: *http://arxiv.org/abs/1904.01226*

Miralinaghi, M. and Peeta, S. (2016), 'Multi-period equilibrium modeling planning framework for tradable credit schemes', *Transportation Research Part E: Logistics and Transportation Review* **93**, 177–198.

Mobility-as-a-Service (n.d.), http://maas.global/what-is-mobility-as-a-service-maas/. Accessed: 2018-06-05.

Nicosia, G., Pacifici, A. and Pferschy, U. (2017), 'Price of fairness for allocating a bounded resource', *European Journal of Operational Research* **257**(3), 933–943.

Nie, Y. (2017*a*), 'Why is license plate rationing not a good transport policy?', *Transportmetrica A: Transport Science* **13**(1), 1–23.

Nie, Y. M. (2012), 'Transaction costs and tradable mobility credits', *Transportation Research Part B: Methodological* **46**(1), 189–203.

Nie, Y. M. (2015), 'A new tradable credit scheme for the morning commute problem', *Networks and Spatial Economics* **15**(3), 719–741.

Nie, Y. M. (2017*b*), 'On the potential remedies for license plate rationing', *Economics of Transportation* **9**, 37–50.

Nie, Y. M. and Yin, Y. (2013), 'Managing rush hour travel choices with tradable credit scheme', *Transportation Research Part B: Methodological* **50**, 1–19.

Rawls, J. (2009), *A theory of justice*, Harvard university press.

Ren, T., Huang, H.-J., Liu, T.-L. and Nie, Y. M. (2020), 'Some analytical results on spatial price differentiation in first–best congestion pricing schemes', *Transportation Research Part C: Emerging Technologies* **114**, 425–445.

Ross, G. T. and Soland, R. M. (1975), 'A branch and bound algorithm for the generalized assignment problem', *Mathematical programming* **8**(1), 91–103.

Rothengatter, T. (1982), 'The effects of police surveillance and law enforcement on driver behaviour', *Current Psychological Reviews* **2**(3), 349–358.

Sakai, K., Liu, R., Kusakabe, T. and Asakura, Y. (2017), 'Pareto-improving social optimal pricing schemes based on bottleneck permits for managing congestion at a merging section', *International Journal of Sustainable Transportation* **11**(10), 737–748.

Schrank, D., Eisele, B. and Lomax, T. (2012), 'Tti's 2012 urban mobility report'.

Schrank, D., Eisele, B., Lomax, T. and Bak, J. (2015), '2015 urban mobility scorecard'.

Shapley, L. S. and Shubik, M. (1971), 'The assignment game i: The core', *International Journal of game theory* **1**(1), 111–130.

Shirmohammadi, N. and Yin, Y. (2016), 'Tradable credit scheme to control bottleneck queue length', *Transportation Research Record: Journal of the Transportation Research Board* (2561), 53–63.

Shirmohammadi, N., Zangui, M., Yin, Y. and Nie, Y. (2013), 'Analysis and design of tradable credit schemes under uncertainty', *Transportation Research Record: Journal of the Transportation Research Board* (2333), 27–36.

Smith, V. L. (2003), 'Constructivist and ecological rationality in economics', *The American Economic Review* **93**(3), 465–508.

Spiess, H. (1990), 'Conical volume-delay functions', *Transportation Science* **24**(2), 153–158.

Su, P. and Park, B. B. (2015), 'Auction-based highway reservation system an agent-based simulation study', *Transportation Research Part C: Emerging Technologies* **60**, 211–226.

Sun, C., Zheng, S. and Wang, R. (2014), 'Restricting driving for better traffic and clearer skies: Did it work in beijing?', *Transport Policy* **32**, 34–41.

Suwansirikul, C., Friesz, T. L. and Tobin, R. L. (1987), 'Equilibrium decomposed optimization: a heuristic for the continuous equilibrium network design problem', *Transportation science* **21**(4), 254–263.

Svensson, L.-G. (1999), 'Strategy-proof allocation of indivisible goods', *Social Choice and Welfare* **16**(4), 557–567.

Tian, L.-J., Yang, H. and Huang, H.-J. (2013), 'Tradable credit schemes for managing bottleneck congestion and modal split with heterogeneous users', *Transportation Research Part E: Logistics and Transportation Review* **54**, 1–13.

Tian, Y. and Chiu, Y.-C. (2015), 'Day-to-day market power and efficiency in tradable mobility credits', *International Journal of Transportation Science and Technology* **4**(3), 209–227.

Trichakis, N. K. (2011), Fairness in operations: from theory to practice, PhD thesis, Massachusetts Institute of Technology.

Tsirimpa, A., Polydoropoulou, A., Pagoni, I. and Tsouros, I. (2019), 'A reward-based instrument for promoting multimodality', *Transportation Research Part F: Traffic Psychology and Behaviour* **65**, 121–140.

Van Vuren, T., Van Vliet, D. and Smith, M. J. (1989), *Combined equalibrium in a network with partial route guidance*, Vol. eds S. Yagar and S. E. Rowe, Engineering Foundation, New York.

Van Vuren, T. and Watling, D. (1991), 'A multiple user class assignment model for route guidance', *Transportation Research Record* pp. 22–22.

Verhoef, E., Nijkamp, P. and Rietveld, P. (1997), 'Tradeable permits: their potential in the regulation of road transport externalities', *Environment and Planning B: Planning and Design* **24**(4), 527–548.

Vickrey, W. (1973), *Pricing, metering, and efficiently using urban transportation facilities*, number 476.

Viegas, J. M. (2001), 'Making urban road pricing acceptable and effective: searching for quality and equity in urban mobility', *Transport Policy* **8**(4), 289–294.

Wada, K. and Akamatsu, T. (2013), 'A hybrid implementation mechanism of tradable network permits system which obviates path enumeration: An auction mechanism with day-to-day capacity control', *Procedia-Social and Behavioral Sciences* **80**, 304–326.

Wada, K., Akamatsu, T. and Nagae, T. (2010), Stochastic convergence of a hybrid scheme of tradable bottleneck permits and congestion pricing, *in* 'Proceedings of the 15th International Conference of Hong Kong Society for Transportation Studies', pp. 173–180.

Wang, G., Gao, Z. and Xu, M. (2019), 'Integrating link-based discrete credit charging scheme into discrete network design problem', *European Journal of Operational Research* **272**(1), 176–187.

Wang, G., Gao, Z., Xu, M. and Sun, H. (2014*a*), 'Joint link-based credit charging and road capacity improvement in continuous network design problem', *Transportation Research Part A: Policy and Practice* **67**, 1–14.

Wang, G., Gao, Z., Xu, M. and Sun, H. (2014*b*), 'Models and a relaxation algorithm for continuous network design problem with a tradable credit scheme and equity constraints', *Computers & Operations Research* **41**, 252–261.

Wang, G., Xu, M., Grant-Muller, S. and Gao, Z. (2018), 'Combination of tradable credit scheme and link capacity improvement to balance economic growth and environmental management in sustainable-oriented transport development: A bi-objective bi-level programming approach', *Transportation Research Part A: Policy and Practice* .

Wang, J.-P., Liu, T.-L. and Huang, H.-J. (2018), 'Tradable od-based travel permits for bi-modal traffic management with heterogeneous users', *Transportation Research Part E: Logistics and Transportation Review* **118**, 589–605.

Wang, J., Peeta, S. and He, X. (2019), 'Multiclass traffic assignment model for mixed traffic flow of human-driven vehicles and connected and autonomous vehicles', *Transportation Research Part B: Methodological* **126**, 139–168.

Wang, P., Wada, K., Akamatsu, T. and Nagae, T. (2018), 'Trading mechanisms for bottleneck permits with multiple purchase opportunities', *Transportation Research Part C: Emerging Technologies* **95**, 414–430.

Wang, X. and Yang, H. (2012), 'Bisection-based trial-and-error implementation of marginal cost pricing and tradable credit scheme', *Transportation Research Part B: Methodological* **46**(9), 1085–1096.

Wang, X., Yang, H., Han, D. and Liu, W. (2014), 'Trial and error method for optimal tradable credit schemes: The network case', *Journal of Advanced Transportation* **48**(6), 685–700.

Wang, X., Yang, H., Zhu, D. and Li, C. (2012), 'Tradable travel credits for congestion management with heterogeneous users', *Transportation Research Part E: Logistics and Transportation Review* **48**(2), 426–437.

Wu, D., Yin, Y., Lawphongpanich, S. and Yang, H. (2012), 'Design of more equitable congestion pricing and tradable credit schemes for multimodal transportation networks', *Transportation Research Part B: Methodological* **46**(9), 1273–1287.

Xiao, F. E., Qian, Z. S. and Zhang, H. M. (2011), 'The morning commute problem with coarse toll and nonidentical commuters', *Networks and Spatial Economics* **11**(2), 343–369.

Xiao, F., Long, J., Li, L., Kou, G. and Nie, Y. (2019), 'Promoting social equity with cyclic tradable credits', *Transportation Research Part B: Methodological* **121**, 56–73.

Xiao, F., Qian, Z. S. and Zhang, H. M. (2013), 'Managing bottleneck congestion with tradable credits', *Transportation Research Part B: Methodological* **56**, 1–14.

Xiao, L.-L., Huang, H.-J. and Liu, R. (2015), 'Tradable credit scheme for rush hour travel choice with heterogeneous commuters', *Advances in Mechanical Engineering* **7**(10), 1687814015612430.

Xiao, L.-L., Liu, T.-L. and Huang, H.-J. (2019), *Transportation* .
URL: *https://doi.org/10.1007/s11116-019-09982-w*

Xu, M. and Grant-Muller, S. (2016*a*), 'Trip mode and travel pattern impacts of a tradable credits scheme: A case study of beijing', *Transport Policy* **47**, 72–83.

Xu, M. and Grant-Muller, S. (2016*b*), 'Vmt reduction and potential environmental effects with a tradable credits scheme: a simulation case study of great britain', *International Journal of Sustainable Development & World Ecology* **23**(6), 514–525.

Xu, M., Mussone, L. and Grant-Muller, S. (2018), 'Effects of a tradable credits scheme on mobility management: A household utility based approach incorporating travel money and travel time budgets', *Case studies on transport policy* **6**(3), 432–438.

Xu, S., Freund, R. M. and Sun, J. (2003), 'Solution methodologies for the smallest enclosing circle problem', *Computational Optimization and Applications* **25**(1-3), 283–292.

Yang, H. (1998), 'Multiple equilibrium behaviors and advanced traveler information systems with endogenous market penetration', *Transportation Research Part B: Methodological* **32**(3), 205–218.

Yang, H. (1999), 'System optimum, stochastic user equilibrium, and optimal link tolls', *Transportation science* **33**(4), 354–360.

Yang, H. and Wang, X. (2011), 'Managing network mobility with tradable credits', *Transportation Research Part B: Methodological* **45**(3), 580–594.

Yang, H., Zhang, X. and Meng, Q. (2007), 'Stackelberg games and multiple equilibrium behaviors on networks', *Transportation Research Part B: Methodological* **41**(8), 841–861.

Ye, H. and Yang, H. (2013), 'Continuous price and flow dynamics of tradable mobility credits', *Transportation Research Part B: Methodological* **57**, 436–450.

Yildirim, M. B. and Hearn, D. W. (2005), 'A first best toll pricing framework for variable demand traffic assignment problems', *Transportation Research Part B: Methodological* **39**(8), 659–678.

Yin, Y. and Yang, H. (2003), 'Simultaneous determination of the equilibrium market penetration and compliance rate of advanced traveler information systems', *Transportation Research Part A: Policy and Practice* **37**(2), 165–181.

Zang, G., Xu, M. and Gao, Z. (2018), 'High-occupancy vehicle lanes and tradable credits scheme for traffic congestion management: A bilevel programming approach', *Promet-Traffic&Transportation* **30**(1), 1–10.

Zhang, K. and Nie, Y. M. (2018), 'Mitigating the impact of selfish routing: An optimal-ratio control scheme (orcs) inspired by autonomous driving', *Transportation Research Part C: Emerging Technologies* **87**, 75–90.

Zhang, X. and Waller, S. T. (2018), 'Implications of link-based equity objectives on transportation network design problem', *Transportation* .
URL: *https://doi.org/10.1007/s11116-018-9888-1*

Zhang, X., Yang, H. and Huang, H.-J. (2011), 'Improving travel efficiency by parking permits distribution and trading', *Transportation Research Part B: Methodological* **45**(7), 1018–1034.

Zhu, D.-L., Yang, H., Li, C.-M. and Wang, X.-L. (2014), 'Properties of the multiclass traffic network equilibria under a tradable credit scheme', *Transportation Science* **49**(3), 519–534.

Zhu, W., Ma, S. and Tian, J. (2017), 'Optimizing congestion and emissions via tradable credit charge and reward scheme without initial credit allocations', *Physica A: Statistical Mechanics and its Applications* **465**, 438–448.

Zhu, X., Wang, F., Chen, C. and Reed, D. D. (2019), 'Personalized incentives for promoting sustainable travel behaviors', *Transportation Research Part C: Emerging Technologies* .
URL: *https://doi.org/10.1016/j.trc.2019.05.015*

Appendix A

In this chapter, we develop some benchmark models, under complete information setting, to evaluate the performance of the proposed MP-based schemes, under incomplete information setting, for single-bottleneck roadways under various scenarios and parameters setting. We construct two different different types of benchmark models, namely, the efficiency-oriented and equity-oriented models.

We assume that the maximum possible utility of mobility user i is finite and limited by a positive value $U_i \in R_{>0}$, which can be converted to monetary terms. By determining the trade-off between their travel time and the unit monetary schedule-early delay and schedule-late delay, users decides on the best departure time to maximize their utility (Doan et al. 2011). In a typical setting, the time-dependent utility of user i, for arriving at time t, can be expressed as: $U_i(t) = U_i - C_i(t)$, $\underline{t} \leq t \leq \bar{t}$, where $C_i(t)$ is their travel cost with departure time t, which can be formulated as

$$C_i(t) = T_i(t)\alpha_i + max\{\beta_i(t_i^* - t - T_i(t)), \gamma_i(t + T_i(t) - t_i^*)\}, \tag{A.1}$$

where $t_i^* \in [\underline{t}, \bar{t}]$ is the preferred arrival time. $T_i(t)$ for $t \in [\underline{t}, \bar{t}]$, is the travel time at departure time t of commuter i. It can also be considered as the queuing delay at the bottleneck, i.e., $T_i(t) = q_i(t)/s$, where $q(t)_i$ is the length of queue the commuter would face at the bottleneck if she departs at time t, and s is the capacity (service rate) of the bottleneck. Parameter $\alpha_i \in R_{>0}$ is the monetary value of unit travel time or VOT of ith commuter. In the same way, $\beta_i, \gamma_i \in R_{>0}$, respectively, are the penalties (costs) for a unit time of early and late arrival of ith commuter. Under this functional form of utility, our hypothetical commuters order their trips in an increasing manner with respect to the values of β_i/γ_i for the early departure and in a decreasing manner with respect to the value of γ_i/β_i for the late departure (Tian et al. 2013). We note that the users' time-dependent net

utility functions characterize their type, which is determined by their $T_i(t), \alpha_i, \beta_i, \gamma_i$, and U_i for each user i. The maximum utility is achieved when a commuter sees no queue at when arriving to the bottleneck. This functional form of users' utilities is a good first-order approximation of the broader class of concave utility functions. By using the lth permit to travel since all queues ($q_i(t) = 0, \forall i$) are eliminated (Liu et al. 2015), the ith commuter will experience a travel cost equal to

$$C_i(t) = max\{\beta_i(t_i^* - tp_l), \gamma_i(tp_l - t_i^*)\} + p_l, \tag{A.2}$$

where p_l is the ask price for l^{th} travel permit. We can assume $t_i^* = q_i^*$ is the most-preferred travel pass expressed by the commuter i, $tp_l = o_l, l = 1, 2, ..., L$, denotes all the available options.

We first develop an efficiency-oriented benchmark model which aims to assign multiple permits to a set of N users ($\{1, 2, ..., N\}$) where each user is guaranteed to get one option. Under symmetric information scenario, we build both the coordinated and centralized system models. To transform the *max* operators in Equation (A.2) to a mathematical programming format, we have defined a binary variable $\tau_{i,l}$ that takes value of one when the lth option is chronologically (time-wise) lower than their most-preferred option, and zero otherwise. The rest of the parameters and variables used in these models are the same as those used in Chapter (4).

A.1 Efficiency-oriented Benchmark Model

$$MPA\text{-}CS: \quad \max \quad \sum_{i=1}^{N}\sum_{k=1}^{L} p_k x_{i,k} \tag{A.3}$$

subject to:

$$\sum_{k=1}^{L} x_{i,k} \quad \leq \quad 1, \quad i = 1,2,...,N, \tag{A.4}$$

$$\sum_{i=1}^{N} o_k x_{i,k} \quad = \quad M_k, \quad k = 1,2,...,L, \tag{A.5}$$

$$o_i^* + BM(1 - \tau_{i,k}) \quad \geq \quad o_k,$$
$$i = 1,2,...,N; k = 1,2,...,L, \tag{A.6}$$

$$o_i^* - BM\tau_{i,k} \quad \leq \quad o_k,$$
$$i = 1,2,...,N; k = 1,2,...,L, \tag{A.7}$$

$$\sum_{k=1}^{L} \tau_{i,k} x_{i,k}(o_i^* - o_k)\beta_i$$
$$+ \sum_{k=1}^{L}(1 - \tau_{i,k})x_{i,k}(o_k - o_i^*)\gamma_i + x_{i,k}p_k \quad \leq$$
$$\tau_{i,l}(o_i^* - o_l)\beta_i$$
$$+ (1 - \tau_{i,l})(o_l - o_i^*)\gamma_i + p_l,$$
$$i = 1,2,...,N; l,k = 1,2,...,L, l \neq k, \tag{A.8}$$

$$U_i \sum_{k=1}^{L} x_{i,k} - \sum_{k=1}^{L}(1 - \tau_{i,k})x_{i,k}(o_i^* - o_k)\beta_i$$
$$- \sum_{k=1}^{L} \tau_{i,k} x_{i,k}(o_k - o_i^*)\gamma_i - \sum_{k=1}^{L} x_{i,k}p_k \quad \geq \quad 0,$$
$$i = 1,2,...,N; k = 1,2,...,L, \tag{A.9}$$

$$x_{i,k}, \tau_{i,k} \quad \in \quad \{0,1\},$$
$$i = 1,2,...,N; l = 1,2,...,L, \tag{A.10}$$

$$p_k \quad \geq \quad 0,$$
$$l = 1,2,...,L. \tag{A.11}$$

In *MPA-CS*, the objective function (A.3) aims to find the optimum decisions (pricing

and allocation of permits) for the efficiency-oriented setting while meeting the constraint set (A.4) to (A.11). The inequality (A.4) states that a commuter should be assigned only to one of available options. In the same way, constraint (A.5) takes care of the capacity limit at each level of identified travel time interval. Using the constraint sets (A.6) and (A.7) we force the variables $\tau_{i,l}$ to take a value of one if the lth option is less (time-wise) than the most-preferred option of ith user, and zero otherwise. Constraint sets (A.8) and (A.9) are the incentive compatibility (IC) and individual rationality (IR) constraints respectively, by which we ensure the participation of the users. To guarantee users' participation, we introduce the IC and IR constraints to the model. In the optimal solution of *MPA-CS* problem, every user is assigned to their top preference (guaranteed by the IC and IR constraints), and the total assigned capacity is less than or equal to the available capacity. We formulated the *MPA-CS* as a nonlinear mixed-integer programming model. In this formulation, we set the $x_{i,l}, i = 1, 2, ..., N; l = 1, 2, ..., L$, as the allocation decision variable, where $x_{i,l}$ takes value one if the commuter $i, i = 1, 2, ..., N$, is allocated with l^{th} option.

The efficiency-oriented system optimum scenario is characterized by the space of all feasible allocations X, and N known utility functions such that $U_i : \boldsymbol{X} \mapsto Z^N_{\geq 0}$ for each user i. Commonly, the system utility $U(x)$ of a solution $x \in X$ is the sum of the users' utilities given by $U(x) = \sum_{i=1}^{N} u_i(x)$, and the optimal solution x^*. The efficiency-oriented system optimum is given by $U^* = U(x^*) = \max_{x \in X} \sum_{i=1}^{N} u_i(x)$ under the side constraints. Given o_l^*, user $i, i = 1, 2, ..., N$, decides to chose option l if and only if $u_i(o_l^*) \geq 0$. Indeed, the *MPA-CS* problem can be readily transformed to the mobility permit allocation problem under system optimum (*MPA-SO*) centralized model by eliminating the pricing decision and considering just the allocation decisions. Given commuters' travel socio-demographic information, the system manager can direct them toward the system-optimal solution state where the travel costs of all travelers combined are minimized (Klein and Ben-Elia 2016). The centralized problem reflects the symmetric information case, which can be modeled by changing the objective function, as simplified in (A.12), and removing the IC, IR, and pricing constraints (A.7), (A.8) and (A.9), respectively.

$$
\begin{aligned}
\textit{MPA-SO:} \quad \max \quad & \sum_{i=1}^{N} U_i \left(\sum_{k=1}^{L} x_{i,k} \right) - \sum_{i=1}^{N} \beta_i \left(\sum_{k=1}^{L} \tau_{i,k} x_{i,k} (o_i^* - o_k) \right) \\
& - \sum_{i=1}^{N} \gamma_i \left(\sum_{k=1}^{L} (1 - \tau_{i,k}) x_{i,k} (o_k - o_i^*) \right) \quad \text{(A.12)} \\
\textit{subject to:} \quad & (A.4), (A.5), (A.6), (A.7), \text{ and } (A.10).
\end{aligned}
$$

A.2 Equity-oriented Benchmark Model

In this section, we develop an equity-oriented benchmark model using maximin fairness principle (Rawls 2009), which is the most acceptable policy pursued by the majority of researchers (Hooker and Williams 2012). The maximin principle guarantees maximum welfare of the worst off user, i.e., $x_{MM} = \arg\max_{x \in X} \min_{i=1,2,...,N} u_i(x)$. This is the only fairness measure that satisfies the max-min fair allocation optimality requirement (Bertsimas et al. 2012). The underlying mobility permit allocation problem under equity-oriented system (*MPA-ES*) can be formulated as follows:

$$MPA\text{-}ES: \quad \max \quad y \tag{A.13}$$

$$\text{subject to:}$$

$$(A.4), (A.5), (A.6), (A.7), \text{ and } (A.10),$$

$$y \quad \leq \quad U_i\left(\sum_{k=1}^{L} x_{i,k}\right) - \beta_i\left(\sum_{k=1}^{L} \tau_{i,k} x_{i,k}(o_i^* - o_k)\right)$$

$$-\gamma_i\left(\sum_{k=1}^{L}(1-\tau_{i,k})x_{i,k}(o_k - o_i^*)\right) \quad i = 1, 2, ..., N. \tag{A.14}$$

We note that, the maximin principle does not guarantee the uniqueness of the obtained solution, as there may be dominated solutions with the same objective function value (Bertsimas et al. 2012).