March 2020

Game Theory Approaches for Transportation Problems

Mahdi Takalloo
University of South Florida

Game Theory Approaches for Transportation Problems

by

Mahdi Takalloo

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Industrial Engineering
Department of Industrial and Management Systems Engineering
College of Engineering
University of South Florida

Major Professor: Changhyun Kwon, Ph.D.
Jamie Kang, Ph.D.
Yu Zhang, Ph.D.
Hadi Charkhgard, Ph.D.
Devashish Das, Ph.D.

Date of Approval:
March 25, 2020

Keywords: Boundedly Rational, Price of Satisficing,
Combinatorial Auction, HazmatToll Pricing, Bi-level Optimization

Dedication

Dedicated to my parents for their endless love and support.

Acknowledgments

I would like to give my biggest appreciation to my adviser Dr. Kwon for his support and patience throughout my PhD. No words can express my gratitude for his dedication and guidance. I am also grateful to the members of my committee, especially Dr. Hadi, for their insightful comments and encouragement through my research. I would like to thank my friends for being an important part of my PhD journey. Last but not least, I am appreciative of my family and my partner for their kindness and support.

Table of Contents

List of Tables

List of Figures

Abstract

This dissertation considers three separate game theory problems in transportation. In the first problem, a combinatorial auction market has been proposed for fractional ownership of autonomous vehicles. The proposed combinatorial auction has two unique features. First, the items are continuous time slots defined by bidders and second, the spatial information of bidders has been incorporated so that sharing becomes a viable plan. A conflict-based formulation of the winner determination problem has been proposed, for which an effective solution approach based on a heuristic and a maximal-clique based relaxation has been presented. The second part of the dissertation examines a pessimistic bilevel toll pricing problem for mitigating the risk of transporting hazardous materials. Since the optimistic hazmat toll pricing problem creates multiple optimal solutions for the inner problem, risk hedging against the behavioral uncertainty of hazmat carriers is desired. Considering hazmat carriers as satisficing decision-makers, an approximate pessimistic problem has been formulated. The optimal solution existence of optimistic and pessimistic problems have been studied. Moreover, solution approaches based on disjunctive programming have been presented. In the third problem, we study the effect of satisficing behavior on transportation network systems. When network users are satisficing decision-makers, the resulting traffic pattern which is called satisficing user equilibrium may deviate from the (perfectly rational) user equilibrium and the total system travel time can be worse than in the case of the perfectly rational user equilibrium. We call the ratio between the total system travel times of the worst-case satisficing user equilibrium traffic pattern and the perfectly rational user equilibrium the price of satisficing, for which we provide an analytical bound. We compare the analytical bound with numerical bounds for several transportation networks.

Chapter 1: Introduction

Game theory is the mathematical modeling of the interactions between a set of players. Each player has a set of choices or strategies and a payoff or utility function, which depends not only on her strategy but also on the strategies taken by other players. Players make choices based on their preferences. In traditional game theory models, it is usually assumed that the players are perfectly rational and have a clear system of preferences and perfect knowledge of the surrounding decision-making environment. As a result, they choose strategies that maximize their outcomes or utilities. Game theory has been widely used in various fields. Researchers have deployed game theory to solve challenging problems in economics, biology, political science, and computer science. In the transportation field, game theory approaches have been used to address different problems, including traffic assignment, matching riders and drivers in ride-sharing, and congestion pricing. In this dissertation, we use game theory approaches to solve three separate transportation problems.

In the first problem, we use auction models to design a new market for fractional ownership autonomous vehicles (AVs). Since the inconveniences of having to locate and walk to a car can be overcome with the self-driving capability, fractional vehicle ownership can become a viable and even cheaper option compare to the car-sharing service or the traditional vehicle ownership. To design a vehicle fractional ownership market, we propose a single-round combinatorial auction that considers spatial information of the bidders. We propose the use of bidder-defined items which enables bidders to bid on any time intervals they wish. In the proposed auction, bidders bid on the package(s) simultaneously. Next, the auctioneer collects all the submitted bids. Considering the spatial information of the bidders, the auctioneer solves the Winner Determination Problem (WDP) to determine the set of winners.

Next, the auctioneer determines the payments based on VCG payment rule. To generate the WDP problem and solve it effectively, we develop a conflict-based formulation. For solving the problem for larger instances, we design a heuristic to find a primal solution and provide a maximal-clique relaxation of the problem to obtain a high-quality dual bound. We compare the performance of the proposed formulations and solution approaches through numerical experiments.

In the second problem, we study an example of Stackelberg games, which is the pessimistic bilevel toll pricing problem for reducing the risk of transporting hazardous materials (hazmat). To reduce the transport risk of hazardous materials, transportation agencies may set tolls on certain road segments to direct the carriers to the less populated and safer roads. One can formulate the hazmat toll pricing problem as a bilevel optimization problem. In the upper-level optimization problem, the transportation agency aims to minimize the transport risk by setting a set of tolls on certain roads. In the inner problem, given a set of road tolls, the network user (shipping carrier) aims to choose the shortest path (the path with the minimum cost). The solution to the inner problem may not be unique. Previous studies have considered an optimistic point of view, which assumes that among the path solutions of the inner problem, the network users choose the one with the lowest risk. However, as we will show through examples and numerical experiments, the solution to the inner problem may not be unique. As a result, the network users may select alternative paths to the optimistic path, which can lead to substantially higher risk than anticipated. With this in mind, we consider a pessimistic point of view and study a setting where the transportation agencies seek a toll vector that minimizes the worst-case risk.

In the third problem, we study the network user equilibrium problem, which is an example of nonatomic games, when network users have an ambiguous system of preferences. When network users are satisficing, the resulting travel pattern is called satisficing user equilibrium (SatUE) and it can be different from a perfectly rational equilibrium (PRUE). We quantify how bad the system travel time in SatUE can be compare to PRUE. In particular, we define the price of satisficing (PoSat) as the ratio between the worst-case total travel time of SatUE and the total travel time of PRUE. We develop analytical bounds on PoSat and compare it to numerical bounds for several networks.

In summary, the objective of this dissertation is to answer the following questions:

- Question 1: How to design a market under which people can co-own a vehicle together?
- Question 2: Why is it important to consider a pessimistic approach for bi-level toll pricing problem and how to formulate and solve pessimistic toll pricing problem?
- Question 3: How bad the total travel time in SatUE can be compare to PRUE when network users are satisficing?

In the process of completing this dissertation and answering the above-mentioned questions, the following papers have been published or submitted:

- Takalloo, M., Bogyrbayeva, A., Charkhgard, H., Kwon, C. Combinatorial Auction with Bidder-Defined Items for Fractional Ownership of Autonomous Vehicles. International Transactions in Operational Research, Under major review.
- Takalloo, M., Huchette, J., Kwon, C. Pessimistic Bi-level Toll Pricing Problem for Mitigating Hazardous Materials Transport Risk. Operations Research, Under review
- Takalloo, M., Kwon, C. On the Price of Satisficing in Network User Equilibria. Transportation Science, Accepted.

The dissertation can be summarized as follows: in Chapter 2, we propose a combinatorial auction for fractional ownership of autonomous vehicles. Chapter 3 presents a pessimistic approach for the bi-level hazmat toll pricing problem to minimize the risk of shipping hazardous materials. In Chapter 4, we discuss How bad the total travel time in SatUE can be compare to PRUE when network users are satisficing. Chapter 5 concludes the dissertation with summarizing of the problems and solution approaches and providing directions for the future studies.

Chapter 2: Combinatorial Auction with Bidder-defined Items for Fractional Ownership of Autonomous Vehicles

2.1 Introduction

People are not interested in owning a car as much as they once were due to several factors, including congestion, parking problems, and the cost of maintaining and operating the vehicle. For this reason, it is no surprise that peer to peer ride-sharing services like Lyft and Uber and car-sharing services like ZipCar have grown significantly in recent years. In response to these recent changes, fractional ownership has also gained attention as a viable alternative to traditional car ownership. By doing so, people can split the cost of leasing. Moreover, given that people have such varying needs, someone who takes two or three trips a week may find fractional ownership even cheaper than a car-sharing service.

The introduction of autonomous vehicles (AVs) to consumer markets will expedite this trend. Even with extreme congestion on the road, "driving" an AV can be a pleasant experience: one can sleep or watch a movie in the car. Moreover, since AVs can travel without passengers, sharing vehicles will become easier. Currently, co-owning a conventional vehicle with friends is not easy, although it can certainly reduce the cost of owning a car. For example, if I want to use a car for commuting and my friend wants to use it while I am at work, my friend must come to my workplace to pick the car up, which requires another form of mobility. With AVs, the co-owned car can travel autonomously from my work to my friend's location. Therefore, we envision that AVs will be co-owned widely.

Co-owning an item is not a new idea. In the private jet market, fractional ownership has been popular (Hicks et al., 2005). One can buy a share in a jet and use it for an allocated amount of time. In the real estate market, especially for resort condominiums, a similar form of ownership exists, known as *timeshare* (Terry, 1994). Multiple owners hold the rights to use a particular property, and each owner is allocated to a certain number of periods of time. The price each owner pays depends on the length of time periods and the season.

Various car manufacturers with conventional vehicles have experimented fractional ownership with some restrictions on the location. Ford Credit Link was a pilot program for co-leasing vehicles in Austin, Texas, but few people registered. Audi had a similar pilot fractional ownership program, 'Audi Unite' in Stockholm, Sweden. After selecting an Audi model and forming a group to co-lease a vehicle of the chosen Audi model. Each customer splits the leasing cost either in proportion to the usage time or at the fixed rate. Nissan Intelligent Get & Go Micra, in Paris, France, used a matching algorithm to form car sharing communities, which partly own a vehicle. All these programs are no longer active.

While the future of this new form of vehicle ownership is unclear with conventional vehicles, it certainly becomes a viable option with an AV. Inconveniences of having to locate and walk to a car can be overcome with the self-driving capability. As typical cars are parked 95% of the time (Shoup, 2005), there is a great potential for AV fractional ownership to change the current structure of car ownership. We note that when people share AVs, *co-leasing* may be a more viable service due to maintenance and insurance issues. In this research, 'fractional ownership' means that customers co-lease a vehicle and 'co-owners' mean co-lessees.

For a practical fractional AV ownership model to be successful, there must be little to no time conflicts among co-owners. Therefore, a suitable mechanism is needed to match customers with non-overlapping time-schedules together and avoid conflicts. With these in mind, in this study, we design an auction market for fractional AV ownership as an alternative to the traditional full ownership model.

Note that although the fractional AV ownership gives less flexibility compared to the dynamic ride-sharing services, it provides a cleaner and cheaper option for the customers who use the vehicle regularly in specific time intervals. Furthermore, customers are subjected to less uncertainty in terms of vehicle availability and the price rates under fractional ownership on these specific time intervals.

2.1.1 Combinatorial Auctions

Combinatorial auctions are suitable mechanisms to sell items or allocate resources in packages, instead of single items separately. Due to the complementarities and substitutability among items, the value of a package of items may not be the same as the summation of the individual values of those items. Furthermore, in many situations, bidders participating in combinatorial auctions have preferences only for a package of items, not for individual items separately. Combinatorial auctions capture these unique characteristics through bidding languages, winner determination methods, and pricing schemes that are specifically designed. For general descriptions of combinatorial auctions, we refer readers to De Vries and Vohra (2003) and Pekeč and Rothkopf (2003).

Combinatorial auctions have been used across various industry sectors. Federal communication commission initiated combinatorial auctions in 1994 to allocate the spectrum right licenses to telecommunication companies. Combinatorial auctions have also been proposed for Internet pricing (Hershberger and Suri, 2001). In transportation and logistics, combinatorial auctions have gained attention for selling airport departure and arrival slots (Rassenti et al., 1982), truckload transportation (Zhang et al., 2015), city bus route market (Cantillon and Pesendorfer, 2006), and railway industry (Kuo and Miller-Hooks, 2015). Recently, researchers suggest combinatorial auctions in ride-sharing market for designing a more efficient shared mobility system (Hara and Hato, 2018), and collaborative vehicle routing (Gansterer and Hartl, 2017) and public transportation systems (James et al., 2018).

In this research, we consider a combinatorial auction market for the fractional AV ownership. In the proposed single-round combinatorial auction, the auctioneer is a car manufacturer or leasing company that wants to sell AVs, and the bidders are customers who are interested in co-leasing a car. The auctioneer sells time-slot packages to bidders through an auction that gives the winners the right to use the same vehicle in these time-slots within a week for a certain period.

Each time-slot package includes several time-slots covering the travel needs of customers. For example, one package can offer the right to use the vehicle for Monday and Tuesday between 8 AM to 9 AM, and Thursday from 6 PM to 7 PM. In the designed auction market, bidders submit multiple bids for various packages, and at most, one of them will be accepted. For instance, if a customer wishes to use the vehicle on Monday between 8 AM to 9 AM, and Tuesday or Thursday from 6 PM to 7 PM, She will submit two bids, one for Monday between 8 AM to 9 AM, and Tuesday from 6 PM to 7 PM, and another one for Monday between 8 AM to 9 AM, and Thursday from 6 PM to 7 PM. Each bid includes a time-slot package and the bidding price offered by the bidder of that package. The bidders must submit their spatial information at the beginning and ending of the time-intervals they intend to use the vehicle. Considering the available vehicles and the spatial information of the bidders, the company determines the set of winning bids to maximize social welfare. Next, the auctioneer determines the payment amount for each winner concerning some pricing rules.

2.1.2 Unique Challenges and Contributions

The setting of the problem under study is unique in several aspects. First, in the most existing combinatorial auctions, products are pre-defined discrete items. In the proposed auction, items are neither pre-defined nor discrete; instead, we consider *bidder-defined continuous-time items*. Every possible time interval is a potential item; it becomes an item when a bidder finds the time interval valuable. As we will discuss later, if we define the items as fixed pre-defined time-slots, we obtain lower social welfare compared to the case where we consider flexible bidder-defined time slots. Second, the fractional vehicle ownership model incorporates the bidder's location as well. If the customers are far away and the gap time between their trips is short, they cannot co-own a car together. A suitable formulation for the model requires not only to consider these unique characteristics but also to provide a suitable framework for designing effective solution approaches as well.

Determining the set of winners of the auction is another challenge. The winner determination problem (WDP) can usually be formulated as a weighted set packing problem, which is known to be NP-hard (Rothkopf et al., 1998). We can categorize methods for solving the WDP into three groups based on the used solution approaches. Given the intractability of finding optimal allocation in combinatorial auctions, the first group of studies (Rothkopf et al., 1998; Tennenholtz, 2000) restricts the set of bids; thereby the problem is solvable in polynomial time. Unfortunately, restricting bids is not possible in many applications, and applying these methods will result in a loss of efficiency. The second group of studies solves the problem to the optimality using various branching methods (Fujishima et al., 1999; Sandholm et al., 2005) The third group of studies develops heuristics approaches, which include virtual simultaneous auctions (Fujishima et al., 1999), equilibrium heuristics (Tsung et al., 2011), and Tabu heuristics (Wu and Hao, 2015).

In this study, considering the spatial information of bidders, we first formulate the WDP in a discrete-time setting where the items are fixed and predefined time slots. In the discrete-time setting, the period under the study is divided into equally-sized time-slots, and bidders can only bid on these predefined time-slots. We use simple examples to show the weakness of the discrete-time formulation. Next, we formulate the problem in a continuous-time setting, where the bidders define items, and we represent its superiority compared to the discrete-time setting. For the continuous-time WDP, we develop a conflict-based formulation, which enables us not only to generate the model more effectively but also to solve it for relatively large-sized instances exactly. To solve large-scale problems, we design the Sequential Single-Vehicle Decomposition (SSVD) heuristic to find a high-quality primal solution. We also develop a maximal-clique based relaxation of the problem to obtain a dual bound for the WDP.

We summarize the contributions of this research as follows. First, this is the first study considering combinatorial auction approaches for the fractional vehicle ownership of AVs. While researchers have studied combinatorial auctions in several areas of transportation, the problem of AV fractional ownership markets has not been addressed yet. Second, we design a new type of combinatorial auctions with unique characteristics, namely, *combinatorial auctions with bidder-defined items*. Instead of using predefined items (time slots in the combinatorial auction market), bidders define them based on their travel needs. Moreover, we incorporate the spatial information of bidders in our formulation and examine whether a package of bids is feasible considering this information. Third, we devise an effective computational method for solving the WDP, which can consider the multiple units and spatial information efficiently. While many desirable properties of the proposed auction market design, based on the well-known Vickery-Clarke-Groves (VCG) mechanism (Vickrey, 1961; Clarke, 1971; Groves, 1973), hold when the WDP is solved optimally, we suggest solving the WDP approximately with a prespecified optimality gap, especially for large-scale problems, as the WDP is NP-hard. To do so, we design the Sequential Single-Vehicle Decomposition (SSVD) heuristic for finding a high-quality primal solution. We also propose a maximal-clique based relaxation to find a dual bound for the WDP. We demonstrate the performance of our solution approach through numerical experiments. Finally, we study the impact of the optimality gap from the WDP on the pricing scheme and the revenue of the auctioneer under VCG payment. We also want to emphasize that we can apply the formulations and the solution approaches proposed here to some other combinatorial auctions when the continuous-time setting and bidder-defined continuous-time items are considered. Such applications may include arrival and departure slots allocation in airports, matching drivers and customers in ride-sharing transportation systems, and resource allocation for cloud computing.

The remainder of this study is organized as follows. In Section 2.2, we discuss some important concepts in combinatorial auctions, including the WDP, incentive compatibility, and rationality. In Section 2.3, we describe the setting of the fractional AV ownership market. We present the WDP under discrete-time and continuous-based settings and compare them in Section 2.4. In Section 2.5, we describe solution approaches to solve the WDP. As mentioned before, the WDP is NP-hard, and it may not be possible to solve it to optimality. We discuss this matter and study its implications on auctioneer's revenue and the payments in Section 2.6. In Section 2.7, we present numerical experiments based on California 2010–2012 travel survey datasets.

2.2 Preliminaries

In this section, we define some fundamental concepts in combinatorial auctions that we use in this study.

2.2.1 The Winner Determination Problem (WDP)

Suppose $\mathcal{I}$ denotes the set of bidders, and set $\mathcal{K}$ represents the set of distinct items. Each bidder $i \in \mathcal{I}$, submits a set of bids $\mathcal{B}_i$. Each bid b_j is a 2-tuple $(c_j, \mathbf{a}_j)$, where c_j is the bidding price for bid j and $\mathbf{a}_j$ is a binary vector of elements a_{kj}, where $a_{kj} = 1$ if item k is in bid j, and is 0 otherwise. We assume that at most one bid from each customer can be determined as a winning bid. The auctioneer pools the submitted bids and determines the winning bids by solving the following WDP:

$$Z^*_{\text{WDP}} = \max_{x_j} \quad \sum_{j \in \mathcal{J}} c_j x_j \tag{2.1}$$

$$\text{s.t.} \quad \sum_{j \in \mathcal{J}} a_{kj} x_j \leq 1 \qquad \forall k \in \mathcal{K} \tag{2.2}$$

$$\sum_{j \in \mathcal{B}_i} x_j \leq 1 \qquad \forall i \in \mathcal{I} \tag{2.3}$$

$$x_j \in \{0,1\} \qquad \forall i \in \mathcal{I}, \forall j \in \mathcal{B}_i \qquad (2.4)$$

which is a special case of the weighted set packing problem. Binary variable x_j specifies whether package j is awarded to the corresponding bidder or not. Objective function (2.1) maximizes the social welfare (the summation of winning bidding prices). Constraint (2.2) ensures that each item assigned to at most one bid. Constraint (2.3) specifies that at most one bid from each bidder can be a winning bid. The optimal solution $\mathbf{x}^*$ to the above problem specifies the optimal allocation.

2.2.2 Incentive Compatibility and Individual Rationality

After determining the set of winners, the auctioneer calculates the payment amount for each winner according to some pricing rules. An ideal payment rule satisfies the two key principles, namely *incentive compatibility* and *individual rationality*. A suitable auction mechanism induces bidders to reveal their true valuations of packages (Pekeč and Rothkopf, 2003). This principle is known as incentive compatibility which ensures that participants in the auction achieve their best outcome by bidding truthfully. The other important principle in auction design is rationality. The bidders will not participate in an auction if they lose money. Hence, the amount of payment for each bidder must be less than or equal to their valuation. A suitable auction design satisfies these two core principles.

2.3 Combinatorial Auction Design for Fractional Ownership of Autonomous Vehicles

In this section, we describe the combinatorial auction setting for the fractional vehicle ownership market.

2.3.1 Items and Bidders

Items are basic units of good or service offered at the auction. In the fractional AV ownership market, the item is the right to use a vehicle at a particular time slot within a specific day in each week for a certain period, which we call a leasing period. For instance, the right to use the vehicle on Monday from 8 PM to 10 PM is a potential item. There are two possible settings for defining the time-slots. In the first setting, the auctioneer defines these time slots, and the bidders can only bid on these time-slots. For instance, the auctioneer may divide each day into the following time slots: 12 AM–6 AM, 6 AM–12 PM, 12 PM–6 PM and 6 PM–12 AM. In this example, the right to use a vehicle on Tuesday between 6 AM–12 PM is a valid item, while the right to use a vehicle on Wednesday between 11 AM–3 PM is not an acceptable item. We call this setting the discrete-time setting. In the second, continuous-time, setting, bidders define time-slots. In a continuous-time setting, bidders can bid on any possible time intervals rather than the restricted discrete fixed predefined time-slots. We will later compare these two settings. We are considering a leasing period from one month to six months for the fractional ownership market. Moreover, we assume a homogeneous vehicle fleet for the proposed auction. If the vehicles are not homogeneous, we can design an auction for each vehicle model separately. Each bidder needs to specify the following information in the submitted bids: *Trip schedule:* Set of items (time intervals) in which bidder wishes to use the vehicle; *Bidding price:* The price offered by the bidder for the corresponding trips; and *Location information:* The spatial information of the vehicle at the origin and the destination of each trip. Spatial information will be used to estimate the commuting time between bidders and consequently find the feasible matches. There is no limit on the number of bids a bidder can submit. However, at most one bid from each customer can be accepted.

2.3.2 The Auction Format

In this study, we use the VCG combinatorial auction, since it is known to achieve both incentive compatibility and rationality. However, we can use the solution approaches for the WDP developed here, in any other mechanism. In this auction, bidders simultaneously submit bids on any individual items or any packages of the items, without having any information about the bids submitted by the others. Next, the auctioneer pools all the bids and solve the WDP by specifying the best allocation of packages to the bidders. Afterward, the auctioneer computes the payments for each winner according to the VCG payment rule, which is as follows: First, bidders express their bids for desirable items or packages of items. Next, the auctioneer solves the WDP and specifies the optimal allocation $\mathbf{x}^*$ and optimal welfare Z^*_{WDP}. Removing winner i's bids from the set of all bids, the auctioneer solves the WDP, denoted by WDP_{-i}, and specifies the optimal allocation $\mathbf{x^*}_{-i}$ and optimal welfare $Z^*_{\text{WDP}_{-i}}$. Then, bidder i pays the amount equal to

$$p_i = Z^*_{\text{WDP}_{-i}} - \left(Z^*_{\text{WDP}} - \sum_{j \in \mathcal{B}_i} c_j x^*_j \right) \tag{2.5}$$

where c_j is the bidding price for bid j. Adding location constraints to ensure the feasibility of the problem do not have any effects on rationality and incentive compatibility of the VCG mechanism if bidders state their locations truthfully. This can be proved by following the proof for incentive compatibility and rationality of general combinatorial auction setting (Ausubel et al., 2006).

Although incentive compatibility and rationality hold under the VCG mechanism, it may be inefficient concerning the auctioneer's revenue (Sandholm and Likhodedov, 2015). Moreover, for large problems, we may not be able to solve the problem to optimality. Under sub-optimal allocation, VCG mechanism does not satisfy compatibility and rationality. We will discuss this matter further in Section 2.6.

2.4 Formulating the Winner Determination Problems

In this section, we formulate the winner determination problem under both the discrete-time setting with predefined time slots and the continuous-time setting with bidder-defined time slots.

2.4.1 The Discrete-Time Winner Determination Problem

Suppose $\mathcal{I}$ denotes the set of bidders participating in the fractional vehicle ownership combinatorial auction. Each bidder $i \in \mathcal{I}$ submits a set of bids $\mathcal{B}_i$, which includes the bidding price c_j, the corresponding time-slots, and the location of the bidder at the origin and destination of each trip. The constant a_{jt} denotes whether a bid j includes a time slot t or not; a_{jt} is 1 if a time-slot $t \in \mathcal{T}$ is included in a bid j and is 0 otherwise. We denote the set of all time slots by $\mathcal{T}$ and the length of each time-slot by Δt. Extracted from bidder's location data, the parameter $r_{iktt'}$ represents the time it takes to commute from bidder i's location at time slot t to a bidder k's location at time slot t' (note that bidders can be in different locations at different times). We formulate the discrete-time WDP as follows:

$$
\begin{aligned}
\text{(P0)} \quad \max_{x_{jv}} \quad & \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{B}_i} \sum_{v \in \mathcal{V}} c_j x_{jv} && && (2.6)\\
\text{s.t.} \quad & \sum_{j \in \mathcal{J}} a_{jt} x_{jv} \leq 1 && \forall t \in \mathcal{T}, v \in \mathcal{V} && (2.7)\\
& \sum_{j \in \mathcal{B}_i} \sum_{v \in \mathcal{V}} x_{jv} \leq 1 && \forall i \in \mathcal{I} && (2.8)\\
& t\Delta t + r_{iktt'} x_{jv} \leq (t'-1)\Delta t + M(1 - x_{lv}) && \forall i, k \in \mathcal{I}, j \in \mathcal{B}_i, l \in \mathcal{B}_k, && (2.9)\\
& && t, t' \in \mathcal{T} : t' > t, a_{jt} a_{lt'} = 1 && \\
& x_{jv} \in \{0, 1\} && \forall i \in \mathcal{I}, j \in \mathcal{B}_i && (2.10)
\end{aligned}
$$

where M is a big positive constant. In the formulation above, the decision variable x_{jv}, is 1 if bid j is assigned to vehicle v and is 0 otherwise. The objective function (2.6) maximizes

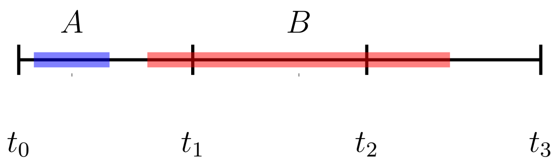


Figure 2.1 – Trip schedules of bidder A and B in Example 2.1

social welfare. Constraint (2.7) ensures that each time slot for each vehicle is assigned to at most one bid. Constraint (2.8) states that at most one bid from each bidder can be determined as a winner. Constraint (2.9) ensures that conflicting bids do not match. As an aside, formulation (P0) can be improved in practice by removing unnecessary time slots from $\mathcal{T}$. In other words, for each trip in each bid, $\mathcal{T}$ can only contain the starting time slot and ending time slot corresponding to that trip. For example, if a given trip in a given bid contains time slots 3–6, then $\mathcal{T}$ should contain time slots 3 and 6 but it is not necessary to insert time slots 4 and 5 in $\mathcal{T}$.

Although the discrete-time WDP provides a suitable framework, it is inferior compared to the continuous-time formulation in terms of social welfare. Moreover, increasing the number of time slots does not necessarily improve the social welfare. The following examples specify the drawbacks of the discrete-time formulation.

Example 2.1. Consider two bidders, bidder A, and bidder B who are bidding in a fractional ownership market with a single vehicle. Suppose the distance between two bidders is negligible. Moreover, assume that the period under study $[t_0, t_3]$ is divided into three time-slots $[t_0, t_1]$, $[t_1, t_2]$, $[t_2, t_3]$. Figure 2.1 represents the trip schedules of bidder A and B. Bidder A bids c_A on time-slot $[t_0, t_1]$, and bidder B bids c_B on time slots $[t_0, t_3]$, $[t_0, t_3]$, $[t_0, t_3]$. Under the discrete-time setting, the social welfare is equal to $\max\{c_A, c_B\}$, while in the continuous-time setting with bidder-defined items, social welfare is equal to $c_A + c_B$ which is strictly greater than $\max\{c_A, c_B\}$ under positive bidding prices.

Example 2.2. Consider the auction in Example 2.1. Suppose bidders A and B have trip schedule represented in Figure 2.2a. Assume that the period under study $[t_0, t_3]$ has been

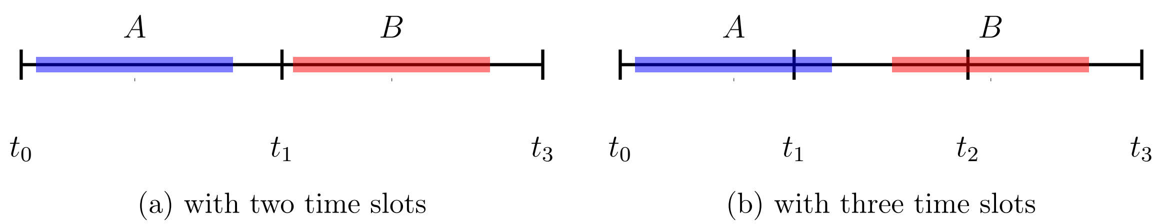


Figure 2.2 – The trip schedules of bidder A and B in Example 2.2

divided into two time-slots $[t_0, t_1]$, $[t_1, t_3]$. Under the discrete-time setting, bidder A bids on the time-slot $[t_0, t_1]$, and bidder 2 bids on the time slot $[t_1, t_3]$. Social welfare is equal to $c_A + c_B$. Now suppose, we increase the number of time-slots from 2 to 3 as depicted in Figure 2.2b. In this case, bidder A bids on time-slot $[t_0, t_1]$ and $[t_1, t_2]$, and bidder B bids on $[t_1, t_2]$ and $[t_2, t_3]$. The social welfare is equal to $\max\{c_A, c_B\}$ which is less than or equal to $c_A + c_B$. Therefore, increasing the number of time-slots may cause even less welfare.

2.4.2 The Continuous-Time Winner Determination Problem

In the continuous-time setting, bidders submit the exact departure and the arrival times of trips along with bidding price and spatial information. Set $\mathcal{T}_j$ denotes the set of trips in bid j. Each trip $n \in \mathcal{T}_j$ is represented by a pair (s_n, e_n), where s_n is the departure time/start time and e_n is the arrival time/end time of that trip. A parameter $r_{ike_ms_n}$ represents the time it takes for an AV to drive from bidder's i location at time e_m to bidder's k location at time s_n. With these notations, we can formulate the WDP under the continuous-time setting as

$$\text{(P1)} \quad \max_{x_{jv}} \quad \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{B}_i} \sum_{v\in\mathcal{V}} c_j x_{jv} \tag{2.11}$$

$$\text{s.t.} \quad \sum_{j\in\mathcal{B}_i} \sum_{v\in\mathcal{V}} x_{jv} \leq 1 \qquad \forall i \in \mathcal{I} \tag{2.12}$$

$$e_m x_{jv} \leq s_n + M(1 - x_{lv}) \qquad \forall i, k \in \mathcal{I}, j \in \mathcal{B}_i, l \in \mathcal{B}_k, \; m \in \mathcal{T}_j, n \in \mathcal{T}_l : s_m \leq s_n \leq e_m \tag{2.13}$$

$$e_m + r_{ike_ms_n} x_{jv} \leq s_n + M(1 - x_{lv}) \qquad \forall i, k \in \mathcal{I}, j \in \mathcal{B}_i, l \in \mathcal{B}_k, \tag{2.14}$$

$$m \in \mathcal{T}_j, n \in \mathcal{T}_l : s_n \geq e_m$$

$$x_{jv} \in \{0,1\} \qquad \forall i \in \mathcal{I}, j \in \mathcal{B}_i \tag{2.15}$$

Constraints (2.13) and (2.14) are equivalent to (2.7) and (2.9) in the discrete-time setting. Constraint (2.13) states that if any two trips from different bids overlap, then those two trips cannot be assigned to a single vehicle. Constraint (2.14) considers the commuting time between the bidder's location and ensures the feasibility of the match.

It is worth to mention that in practice the start time and end time of trips are subject to some uncertainty. For instance, bidder k may have some delay and finish her trip at time $e'_m > e_m$. As a result, we may not be able to serve bidder i at the pre-scheduled time s_n. To deal with this uncertainty, the auctioneer can set $r_{ike_ms_n}$ to a value that is greater than the commuting time between bidders' locations by some margin.

Problem (P1) is a binary problem with many constraints and variables. Generating the optimization model for (P1), specially Constraint (2.13) and Constraint (2.14) is time consuming. Moreover, these two constraints include big-M constraints, which make the problem more difficult to solve. In the next section, we propose a conflict-based approach to eliminate these constraints and construct and solve the model faster.

2.5 Computational Methods

In this section, we introduce an equivalent conflict-based formulation for the continuous-time WDP. To solve the conflict-based problem, we need to design an algorithm that finds the conflicting bids efficiently. We introduce such an algorithm in Section 2.5.2. We introduce SSVD algorithm to obtain a high-quality feasible solution for the WDP in short time. We also introduce the so-called maximal-clique based relaxation of the problem to find a dual bound for the problem.

Table 2.1 – Time-slots and bidding prices for submitted packages in Example 2.3

Package	Time Intervals			Bidding Price
1	6:00 AM–9:00 AM			45
2	6:00 AM–6:30 AM	2:30 PM–3:30 PM		25
3	8:00 AM–8:30 AM	3:00 PM–4:00 PM		20
4	2:00 PM–4:00 PM			20
5	10:00 AM–12:00 PM	2:00 PM–4:00 PM	6:00 PM–7:30 PM	55

2.5.1 The Conflict-based Reformulation of the WDP

We can rewrite (P1) in an equivalent conflict-based formulation. Consider graph $G = (\mathcal{N}, \mathcal{A})$, in which $\mathcal{N}$ represents the set of nodes and set $\mathcal{A}$ represents the set of arcs. Each node $n_j \in \mathcal{N}$ represents a bid and has a weight which is equal to the corresponding bidding price c_j. Each edge connects two nodes with overlapping bids.

Example 2.3. Suppose there are five bidders in a fractional vehicle ownership combinatorial auction. Table 2.1 shows the time-slots associated with each bidder's bid as well as the bidding price for those packages. Suppose that the commuting times between bidder's locations is negligible. The set of all non-matchable pairs is

$$\big\{\{1,2\},\{1,3\},\{2,3\},\{2,4\},\{2,5\},\{3,4\},\{3,5\},\{4,5\}\big\}.$$

We can have the following conflict-based constraints for this problem:

$$\begin{aligned}
&x_1 + x_2 \le 1, \quad x_1 + x_3 \le 1, \quad x_2 + x_3 \le 1, \quad x_2 + x_4 \le 1,\\
&x_2 + x_5 \le 1, \quad x_3 + x_4 \le 1, \quad x_3 + x_5 \le 1, \quad x_4 + x_5 \le 1,\\
&x_i \in \{0,1\} \quad \forall i \in \{1,2,3,4,5\}.
\end{aligned}$$

In the light of Example 2.3, Problem (P1) can simply be written as follows:

$$\text{(P2)} \quad \max_{x_{jv}} \quad \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{B}_i} \sum_{v\in\mathcal{V}} c_j x_{jv} \tag{2.16}$$

$$\text{s.t.} \quad \sum_{j\in\mathcal{B}_i}\sum_{v\in\mathcal{V}} x_{jv} \leq 1 \quad \forall i \in \mathcal{I} \tag{2.17}$$

$$x_{jv} + x_{lv} \leq 1 \quad \forall v \in \mathcal{V}, i, k \in \mathcal{I}, j \in \mathcal{B}_i, l \in \mathcal{B}_k : j, l \text{ are conflicting} \tag{2.18}$$

$$x_{jv} \in \{0, 1\} \quad \forall j \in \mathcal{B}, \forall v \in \mathcal{V} \tag{2.19}$$

Constraint (2.18) is a conflict constraint, which ensures that two conflicting bids do not share a vehicle. Two bids do not match if they have at least one overlapping intervals. Mathematically speaking, bid $b_j \in \mathcal{B}_i$ *conflicts* with bid $b_l \in \mathcal{B}_k$ if and only if

$$s_m \leq e_n + r_{ike_ns_m}, \quad s_n \leq e_m + r_{ike_ms_n} \quad \text{for some } m \in \mathcal{T}_j, n \in \mathcal{T}_l \tag{2.20}$$

Constraint (2.18) can be easily generated in relatively short time compared to (2.13)–(2.14). Moreover, it makes the problem easier to solve. Next, we prove that formulations (P2) and (P1) are equivalent. Proofs for all propositions are provided in the appendix.

Proposition 2.1. (P2) is equivalent to (P1).

To use formulation (P2), we need to find the set of conflicting bids, for which we present a simple algorithm in the next section.

2.5.2 Detecting Conflicting Bids

As described above, formulation (P2) requires finding all overlapping bids. Each bid includes several trips. Two bids conflict if they have overlapping trips. Algorithm 1 represents an exact algorithm to find the set of conflicting bids in polynomial time.

Proposition 2.2. Let p be the total number of bids in a fractional vehicle ownership combinatorial auction, i.e., $p = |\mathcal{B}|$. Also, let q be the maximum number of trips in a bidding package. The time complexity of Algorithm 1 is $O(p^2 q \log q)$.

Once we determine the conflicting bids, we can solve problem (P2) by CPLEX or any other integer programming solver.

Algorithm 1: Detecting conflicting bids

```
Input: Start time, end time, and location information of the trips in all bids
Output: Set of conflicting bids
foreach pair of bids belonging to bidders i and k do
    Place the start and end times of all the trips in the pair (of bids) in vector L
      such that the end time of each trip is right after its start time. Also, make sure
      that all the even cells contain the end times and all the odd cells contain the
      start times;
    Create a binary vector I with the same size as vector L;
    Set the value of each cell in I to i if its corresponding value in vector L belongs
      to the bid submitted by bidder i, and set it to k otherwise;
    Sort vector L in a non-decreasing order and update I accordingly;
    t ← 2;
    while t < length(L) do
        if I[t] ≠ I[t − 1] then
            The pair is conflicting, report it, and go to the next pair;
        end
        if I[t] ≠ I[t + 1] then
            if L[t + 1] ≤ L[t] + r_{I[t]I[t+1]L[t]L[t+1]} then
                The pair is conflicting, report it, and go to the next pair;
            end
        end
        t ← t + 2
    end
end
```

2.5.3 Sequential Single-Vehicle Decomposition (SSVD) Heuristic

It usually takes a considerable amount of time to find a high-quality feasible solution for problem (P2) for large instances. In this section, we introduce SSVD to find a high-quality solution in a relatively short time.

Since we assume a homogeneous fleet of vehicles, we can decompose the vehicle fractional ownership combinatorial auction to a $|\mathcal{V}|$-round single vehicle combinatorial auction. At each round, considering the set of remaining bidders, we solve the WDP for a single vehicle and find the winners. Then, we update the set of bidders by excluding the winners from the list of bidders and go to the next round. This procedure continues until we assign all the vehicles to the bidders. The SSVD algorithm steps has been presented in the following pseudo-code.

Algorithm 2: SSVD heuristics for computing a high-quality feasible solution

Input: Set of vehicles $\mathcal{V}$, set of bidders $\mathcal{I}$, and set of bids $\mathcal{B}_i$ for each $i \in \mathcal{I}$,
Output: Set of winners $\mathcal{W}$

```
Initialization W ← ∅;
for v = 1 to |V| do
    Consider an auction with a single vehicle v;
    Solve problem (P2) with the set of bidders I;
    Ŵ ← the set of winners of the auction ;
    W ← W ∪ Ŵ;
    I ← I \ Ŵ;
end
```

We can bound the quality of the solution found by SSVD as the following proposition presents.

Proposition 2.3. Given the same set of bids and corresponding bid prices, let $Z^*_{\mathcal{V}}$ and $Z^G_{\mathcal{V}}$ be the optimal objective function value for (P2) and the SSVD objective value when the set of vehicles is $\mathcal{V}$. Let $Z^*_{\mathcal{V}}$ be the optimal objective value when the set of vehicles is $\mathcal{V}'$, where $|\mathcal{V}| > |\mathcal{V}'|$. Then we have

$$Z^*_{\mathcal{V}} - Z^G_{\mathcal{V}} \leq \frac{|\mathcal{V}|}{|\mathcal{V}'|} Z^*_{\mathcal{V}'} - Z^G_{\mathcal{V}}. \tag{2.21}$$

When the fleet size ($|\mathcal{V}|$) is large, we fail to solve (P2) to optimality. However, we can solve (P2) optimally when the fleet size is relatively small. Proposition 2.3 utilizes this fact to find an optimality gap bound for the solutions found by SSVD. In the first round, SSVD will find $Z^*_{\{v_1\}}$, with v_1 being the first vehicle the algorithm considers; this information can bound the performance of the algorithm as follows:

$$Z^*_{\mathcal{V}} - Z^G_{\mathcal{V}} \leq |\mathcal{V}| Z^*_{\{v_1\}} - Z^G_{\mathcal{V}}.$$

In general, however, the bound in (2.21) does not provide a high-quality dual bound. In Section 2.5.4, we introduce a relaxation of (P2), which is based on finding maximal cliques, to find a better dual bound for the problem.

Note that problem (P2) has symmetry issues. An integer linear program is symmetric if its variables can be rearranged without structural changes in the problem (Margot, 2010). To see this, consider Example 2.3 and suppose the fleet size for the auction is 2. One possible solution is to assign bids $\{1, 5\}$ to vehicle 1, and $\{2\}$ to vehicle 2. However, we obtain the same objective value by assigning $\{1, 5\}$ to vehicle 2 and $\{2\}$ to vehicle 1. This symmetry problem, which arises from constraint (2.17), makes the problem difficult to solve as the fleet size increases. Since SSVD solves the WDP with a single vehicle in each iteration, the symmetry issue is eliminated. In Section , we develop a relaxation for (P2), which breaks the symmetry.

2.5.4 The Maximal-Clique Based Relaxation for Finding Dual Bounds

To evaluate the quality of the solution obtained from SSVD, it is necessary to find a good dual bound. Computing a good dual bound can be done by developing a good relaxation. A good relaxation can, in turn, be obtained based on good formulation. So, in this section, we first show that how a good formulation can be obtained and then we discuss a how a good relaxation can be developed.

Note that each conflict constraint (2.18) can be viewed as a simple *clique* constraint. In graph theory a clique is a complete subgraph of a given graph. By this definition, each conflict constraint (2.18) is simply constructed based on a clique with cardinality of two in the conflicting graph. This observation enables us to derive a better formulation by replacing constraints of the form (2.18) with stronger constraints based on cliques with larger cardinalities. To this end, we can use maximal cliques. Note that a maximal clique is a clique that cannot be extended by adding any other node. In Example 3, we can find two maximal cliques, namely $\{1, 2, 3\}$ and $\{2, 3, 4, 5\}$. Accordingly, we can replace conflict constraints in Example 2.3 by the following stronger constraints:

$$x_1 + x_2 + x_4 \leq 1, \qquad x_2 + x_3 + x_4 + x_5 \leq 1.$$

These two constraints give us a stronger formulation. For instance, $\mathbf{x} = (0.5, 0.5, 0, 0.5, 0)$ does not satisfy the above constraints, while it satisfies constraint (2.18).

In light of the above, we can obtain a stronger formulation by using maximal cliques:

$$\text{(P3)} \quad \max_{x_{jv}} \quad \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{B}_i} \sum_{v \in \mathcal{V}} c_j x_{jv} \tag{2.22}$$

$$\text{s.t.} \quad \sum_{j \in \mathcal{B}_i} \sum_{v \in \mathcal{V}} x_{jv} \leq 1 \quad \forall i \in \mathcal{I} \tag{2.23}$$

$$\sum_{j \in \mathcal{C}_m} x_{jv} \leq 1 \quad \forall v \in \mathcal{V}, m \in \mathcal{M} \tag{2.24}$$

$$x_{jv} \in \{0, 1\} \quad \forall i \in \mathcal{I}, j \in \mathcal{B}_i, \forall v \in \mathcal{V} \tag{2.25}$$

where $\mathcal{M}$ is the set of all maximal cliques, and $\mathcal{C}_m$ is the set of nodes in maximal clique $m \in \mathcal{M}$. The LP relaxation of (P3) is superior to the LP relaxation of (P2). Commercial solvers such as CPLEX add similar cliques to the problem in the preprocessing phase to tighten LP feasible region of the problem.

Based on (P3), we now develop a relaxation problem to compute dual bounds. In order to do so, we use an aggregation technique to reduce the number of decision variables. Note that problem (P3) still suffers from symmetry issue because of having homogeneous vehicles. So, using an aggregation technique over all vehicles we can not only reduce the size of the problem but also break the symmetry. Specifically, we define $y_j := \sum_{v \in \mathcal{V}} x_{jv}$ for each $j \in \mathcal{B}$ and introduce the following relaxation of problem (P3):

$$\text{(R)} \quad \max_{y_j} \quad \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{B}_i} c_j y_j \tag{2.26}$$

$$\text{s.t.} \quad \sum_{j \in \mathcal{B}_i} y_j \leq 1 \quad \forall i \in \mathcal{I} \tag{2.27}$$

$$\sum_{j \in \mathcal{C}_m} y_j \leq |\mathcal{V}| \quad \forall m \in \mathcal{M} \tag{2.28}$$

$$y_j \in \{0, 1\} \quad \forall i \in \mathcal{I}, j \in \mathcal{B}_i \tag{2.29}$$

where constraint (2.28) states that at most $|\mathcal{V}|$ bids can be determined as a winner among a group of conflicting bids, since there are only $|\mathcal{V}|$ available vehicles.

Proposition 2.4. Problem (R) is a relaxation of problems (P2) and (P1).

Now, one can solve problem (R) to compute a dual bound. It is evident that this problem is expected to be solve quickly because of breaking symmetry and having less decision variables. However, a key issue is that it is not always possible to add all maximal cliques to the problem because (1) finding all maximal cliques is NP-hard (Östergård, 2002), and (2) the number of maximal cliques increases exponentially. Therefore, we choose a subset of maximal cliques randomly. We use the following procedure to generate maximal cliques: We start to create a clique by adding one random bid to it. Next, we attempt to add all the other bids in a random order, if possible. We add a bid to the clique if it conflicts with all the existing bids in the clique. We repeat this process multiple times and impose a time limit as the termination condition. Also, we make sure that all the bids will initiate a clique at least once in our procedure. As we will show in Section 2.7, problem (P3) gives us a better dual bound compare to the CPLEX dual bound for problem (P2) for large instances.

2.6 VCG Mechanism under Suboptimal Solutions

It usually takes a considerable amount of time to solve the WDP optimally. However, it is possible to solve relatively large problems with a small optimality gap. When the solution is suboptimal, incentive compatibility and rationality do not necessarily hold under the VCG mechanism. Moreover, suboptimal allocation affects the winners' payments and consequently the revenue of the auctioneer. In the following, we study VCG mechanism under suboptimal solution.

Incentive compatibility does not necessarily hold under suboptimal allocation algorithms (Nisan and Ronen, 2007). Although in theory suboptimal allocation may influence the bidders to mistrust, in practice, suitable algorithms essentially eliminate agent's incentives to misreport their preferences (Vorobeychik and Engel, 2011). Next, we study the effect of suboptimality on rationality assumption. Let Z^{ϵ}, $\mathbf{x}^{\epsilon}$ and $\mathbf{p}^{\epsilon}$ denote the social welfare, the package allocation and the bidder's payment when the WDP is solved with the optimality gap of ϵ, respectively. That is,

$$\frac{Z^{*}_{\text{WDP}} - Z^{\epsilon}_{\text{WDP}}}{Z^{\epsilon}_{\text{WDP}}} \leq \epsilon, \tag{2.30}$$

The rationality assumption holds if $v_i^{\epsilon}(\mathbf{x}^{i,\epsilon}) \geq p_i^{\epsilon}$ for each $i \in \mathcal{I}$, where v_i^{ϵ} is the bidder's i true valuation for the packages awarded to her, and $\mathbf{x}^{i,\epsilon}$ is the package allocation for the bids submitted by bidder i. If bidders bid truthfully, under certain conditions, the rationality assumption holds.

Proposition 2.5. Suppose ϵ represents the optimality gap for the winner determination problem and the incentive compatibility assumption holds. Then the rationality assumption holds if $Z^{\epsilon}_{\text{WDP}} \geq Z^{\epsilon}_{\text{WDP}_{-i}}$ for all $i \in \mathcal{I}$.

Note that when $\epsilon = 0$, we have $Z^{0}_{\text{WDP}} = Z^{*}_{\text{WDP}} \geq Z^{*}_{\text{WDP}_{-i}} = Z^{0}_{\text{WDP}_{-i}}$ since the problem is solved optimally. However, in general, when $\epsilon > 0$, inequality $Z^{\epsilon}_{\text{WDP}} \geq Z^{\epsilon}_{\text{WDP}_{-i}}$ is not necessarily true. Nevertheless, if bidders bid truthfully, and removing any bidder does not increase the social welfare under suboptimal allocation, the rationality holds.

2.6.1 Bidder's Payments and Auctioneer's Revenue

A suboptimal allocation affects the winners' payments and the revenue of an auctioneer. The following proposition compares the payments and the auctioneer's revenue under an optimal and suboptimal allocation.

Proposition 2.6. Let p_i^* and p_i^ϵ denote the payments of a bidder i under the optimality and a optimality gap of ϵ, and R^* and R^ϵ denote the auctioneer's revenue under the optimality and an optimality gap of ϵ under the VCG mechanism, respectively. Then the following bounds hold:

$$|p_i^\epsilon - p_i^*| \le \epsilon \max\left\{Z^\epsilon_{\text{WDP}}, Z^\epsilon_{\text{WDP}_{-i}}\right\} + \max_{j \in \mathcal{B}_i} c_j, \tag{2.31}$$

$$|R^\epsilon - R^*| \le \epsilon \max\left\{(|\mathcal{I}| - 1)Z^\epsilon_{\text{WDP}}, \sum_{i \in \mathcal{I}} Z^\epsilon_{\text{WDP}_{-i}}\right\}. \tag{2.32}$$

As ϵ approaches to zero, the upper bound (2.32) on the revenues difference goes to 0. The bound (2.31) on the payments difference, however, goes to $\max_{j \in \mathcal{B}_i} c_j$; therefore, it is not tight. The revenue and a bidder's payments under a suboptimal allocation can be higher or smaller than the revenue and the bidder's payments under the optimal allocation.

2.7 Numerical Experiments

We compare the discrete model and the continuous model numerically and evaluate the performance of (P1) and (P2). We use the Julia 7.0 to implement all formulations and algorithms. We also use CPLEX 12.5 to solve integer programming problems.

To test our results, we use 2010–2012 California Household Travel Survey[1], which includes the travel information of 2908 vehicles in a week. The survey dataset includes the date, departure time, arrival time, duration, mileage, and origin and destination of each trip for every vehicle under the study. We test the performance of the proposed formulations

[1] https://trid.trb.org/view/1308918

and algorithms by generating some instances using 2010–2012 California Household Travel Survey. To generate the different bids other than the travel pattern for each bidder, we assume that bidders drop one or two trips randomly. To generate the bidding prices for a leasing period of four weeks, we use the following formula:

$$c_b = 4 \times (5 + 0.76\, \widetilde{a}_b m_b)$$

where we assume the fixed leasing cost for each week is \$5 and the average cost per mile is \$0.76 [2], and $\widetilde{a}_b$ is randomly drawn from the interval $[0.8, 1.2]$ and m_b is the total mileage of all the trips in the package. The location distances of each pair of customers are randomly selected between 0 and 60 minutes. The number of bidders and the fleet size in the experiments vary, depending on the classes of instances used. We will describe these classes when we present the numerical results for each experiment.

Figure 2.3 compares the social welfare under discrete-time problem (P0) and the continuous time problem (P1) for a simulated fractional ownership auction with a single vehicle and 100 bidders. As presented, the continuous-time model is always superior to discrete-time model with respect to social welfare. Moreover, increasing the number of time slots in the discrete-time setting does not necessarily improve social welfare as it has been explained in Example 2.2.

We compare the model building time (the time that Julia takes to generate the model) and the solution time of (P1) and (P2) to show the superiority of (P2). For the comparison, we generate ten classes of instances based on the fleet size and the number of bids. We set the number of bidders to 100, 150, 200, 250 and 300, and the number of vehicles to 1 and 5. We generated five instances for each class, and use the performance profile to compare the building time and the solution time of the basic formulation (P1) and the conflict-based formulation (P2). We construct the performance profile in Figure 2.4a for comparing model building time as follows: For each formulation and for each problem instance, we compute the

[2] American Automobile Association, 2015 report

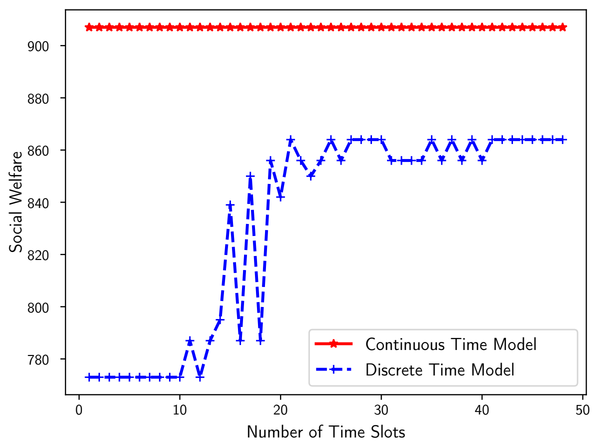
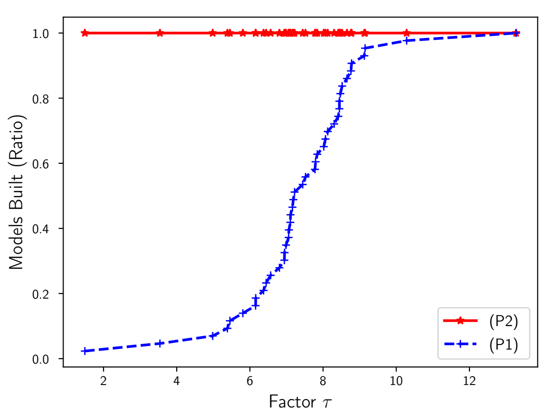
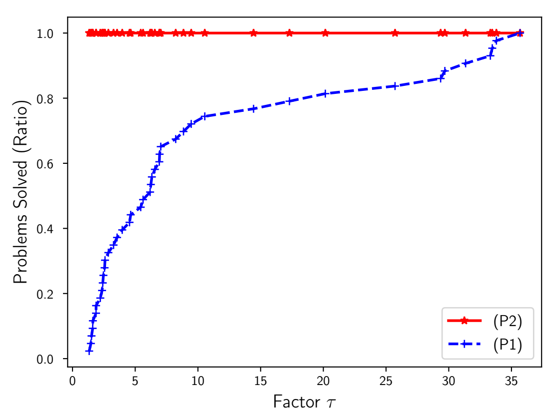


Figure 2.3 – Comparing social welfare under the continuous-time setting and the discrete-time setting for simulated fractional ownership auction



(a) Model Building Time



(b) Model Computation Time

Figure 2.4 – Comparing the basic formulation (P1) and the conflict-based formulation (P2) for simulated fractional ownership auction

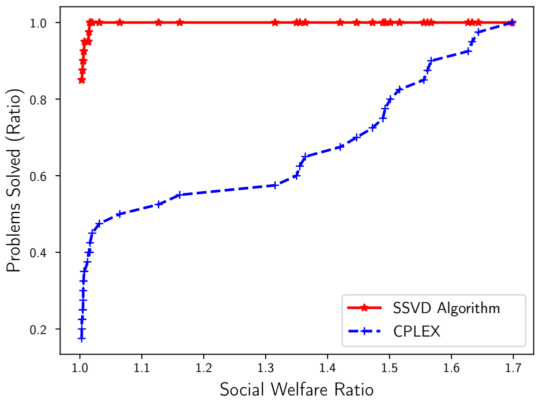
Table 2.2 – Computational time and optimality gap for simulated fractional ownership auction for small instances

$\lvert\mathcal{I}\rvert$	$\lvert\mathcal{B}\rvert$	$\lvert\mathcal{H}\rvert$	R	$\mathrm{T}^{\mathrm{SSVD}}$(sec)	$\mathrm{T}^{\mathrm{CPLEX}}$(sec)	$\mathrm{Gap}^{\mathrm{SSVD}}$	$\mathrm{Gap}^{\mathrm{CPLEX}}$
100	292.6	5	1	2.41	6.47	0.83	0
150	438.8	5	1	8.30	32.72	0.87	0
200	582.8	5	1	20.28	120.77	1.58	0
250	730	5	1	37.79	535.22	1.37	0
300	876.2	5	0.4	55.62	1526.69	4.09	1.93
100	292.6	10	0.9	3.28	162.47	1.49	0.04
150	438.8	10	0.3	12.41	275.44	3.20	1.35
200	582.8	10	0.1	31.75	3379.70	5.03	2.93
250	730	10	0	64.86	> 3600	6.95	5.76
300	876.2	10	0	100.39	> 3600	13.26	11.53

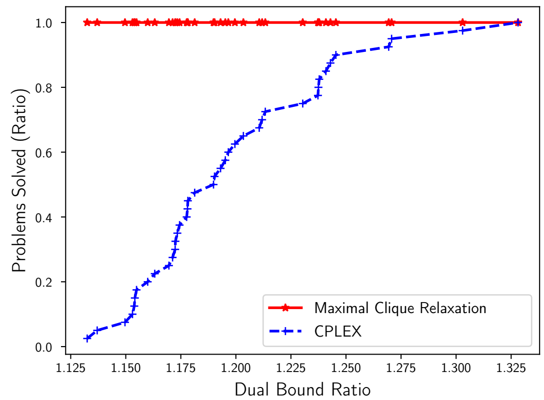
ratio between the model building time of the formulation on the instance and the minimum of the model building time of two formulations on the instance. The performance profile in Figure 2.4a then shows for each formulation, on the vertical axis, the fraction of instances with a ratio that is less than or equal to the factor τ shown on the horizontal axis. The performance profile in Figure 2.4b for solution time can be created similarly.

For almost all instances, the model building time and the solution time for the conflict-based formulation (P2) is smaller than the basic continuous formulation (P1). We could generate the conflict-based model 13 times faster compared to the basic continuous model in (P1). Moreover, we could solve (P1) 35 times faster than (P2). Formulation (P1) is difficult to solve since it has big-M constraints. However, the conflict-based formulation has only clique constraints. CPLEX utilizes these clique constraints to generate stronger constraints and solves the model much faster.

We test the performance of the conflict-based formulation by solving (P2) for different instances. Table 2.2 summarizes our results for smaller instances. We consider ten classes of instances based on the number of bidders and fleet size, as can be seen in Table 2.2, and generate ten instances for each class. We report the SSVD average computation time $\mathrm{T}^{\mathrm{SSVD}}$, the CPLEX average computation time for the problems solved to optimality within one hour $\mathrm{T}^{\mathrm{CPLEX}}$, and the ratio of instances in each class solved to optimality (R) by CPLEX within



(a) Primal Solution Ratio



(b) Dual Bound Ratio

Figure 2.5 – Comparing primal solution and dual solution with CPLEX solution for simulated fractional ownership auction

an hour , the gap between the primal bound obtained by CPLEX and SSVD and the CPLEX dual bound. When the number of vehicles is relatively small ($|\mathcal{H}| = 5$), we could solve 44 out of 50 instances to optimality within an hour by CPLEX, however, when the number of vehicles is larger ($|\mathcal{H}| = 10$), the symmetry becomes more severe, and CPLEX fails to solve the problem to optimality. However, SSVD can find a primal solution within 2 minutes. Moreover, as presented in Table 2.2, the gap between the primal bound obtained by CPLEX and SSVD is negligible (within 2.2%), which suggests that SSVD provides a high-quality primal solution. The optimality gap of SSVD is below 10% for all classes of instances except the last class, which has the highest average number of bids ($|\mathcal{B}| = 876.2$) and number of vehicles ($|\mathcal{H}| = 10$).

Next, we compare the performance of the SSVD and the maximal clique relaxation with CPLEX for larger instances. For comparison purposes, we consider four classes of instances based on the fleet size. We set the fleet size to 15, 20, 25 and 30. We set the number of bidders to 300 for all classes. We generate ten instances for each class, and compare the solution of the SSVD heuristics with CPLEX through the performance profile presented in Figure 2.5a. We set the CPLEX solution time limit to one hour. As Figure 2.5a shows, the SSVD heuristics outperforms in 33 out of 40 instances. The solution found by the SSVD

Table 2.3 – Optimality gap for simulated fractional ownership auction for large instances

$\lvert\mathcal{I}\rvert$	$\lvert\mathcal{B}\rvert$	$\lvert\mathcal{H}\rvert$	$\mathrm{T}^{\mathrm{SSVD}}$(sec)	$\mathrm{T}^{\mathrm{CPLEX}}$(sec)	$\frac{Z^{\mathrm{SSVD}}}{Z^{\mathrm{CPLEX}}}$	$\mathrm{Gap}^{\mathrm{SSVD}}$
300	879.2	15	147.77	>3600	1.00	15.00
300	879.2	20	178.08	>3600	1.10	13.97
300	879.2	25	200.32	>3600	1.36	12.74
300	879.2	30	217.48	>3600	1.38	11.49

heuristics can be as high as 1.7 times of the CPLEX solution. Figure 2.5b compares the dual bound found by (R) with CPLEX dual bound, found within an hour, For all instances, maximal clique relaxation (R) outperforms CPLEX dual bound. The dual bound found by (R) can outperform CPLEX by more than 30 %.

Table 2.3 reports the average solution time for each class, and the optimality gap for large instances obtained by comparing SSVD primal solution and maximal clique dual solution. It also reports the SSVD to CPLEX objective value ratio ($\frac{Z^{\mathrm{SSVD}}}{Z^{\mathrm{CPLEX}}}$). As it can be seen in the table, the SSVD primal bound outperforms CPLEX. Moreover SSVD is much faster and can find the primal bound in less than 4 minutes for all classes of instances. As seen in Table 2.3, while the optimality gap is considerable for larger instances (from 11.49% to 15%), the SSVD primal bound outperforms CPLEX.

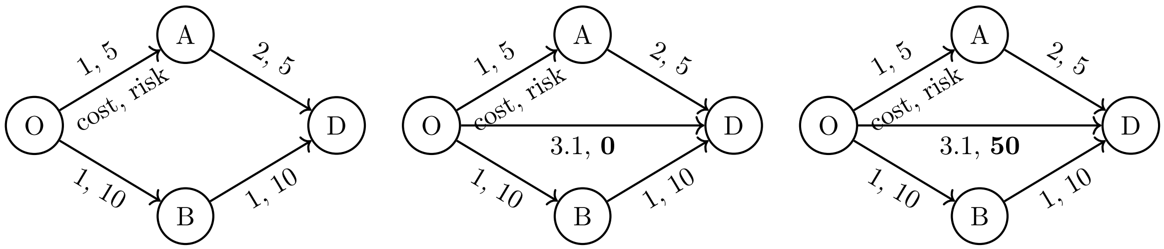
Chapter 3: Pessimistic Bilevel Toll Pricing Problem for Mitigating Hazardous Materials Transport Risk

3.1 Introduction

Hazardous materials (hazmat) are items and chemicals classified by the United States Department of Transportation as potentially dangerous to the safety of the environment or population. The general public could potentially encounter risk when an accident involving a shipment of hazmat occurs. Therefore, specific policies and regulations are needed for shipping hazardous materials. To protect against the risks of transportation of hazardous material, the transport of hazmat is regulated under the Federal Hazardous Materials Transportation Act under the authority of the United States Secretary of Transportation.

Two main policies have been proposed for reducing the hazmat transport risks: network design and toll-setting. Network design policy (Kara and Verter, 2004), which has been practiced by transportation agencies, forbids the use of particular road segments for some hazmat carriers. However, such a policy is considered as being too inflexible, since it completely bans some shipping carriers from using certain roads. Alternatively, a toll policy involves setting tolls on certain road segments to channel the carriers to the less populated roads. It is more flexible and effective than the network design policy for reducing transport risk (Marcotte et al., 2009). In this study, we consider the toll-setting policy for minimizing the exposure risk of hazmat.

One can formulate the hazmat toll pricing problem as a bilevel optimization problem. In the upper-level optimization problem, the transportation agency aims to minimize the



(a) Two paths from O to D (b) A new path with low risk (c) A new path with high risk

Figure 3.1 – Simple network for comparing the optimistic and the pessimistic approach for hazmat toll pricing problem

transport risk by setting a set of tolls on certain links. In the lower level or inner problem, given a set of road tolls, the network user (shipping carrier) aims to choose the shortest path (the path with the minimum cost). The solution to the inner problem may not be unique. Previous studies have considered an optimistic point of view, which assumes that among the path solutions to the inner problem, the network users choose the one with the lowest risk. However, as we will show through examples and numerical experiments, the solution to the inner problem under the toll-setting policy may not be unique. As a result, the network users are able to select alternatives to the "optimistic" path, which can lead to substantially higher risk than anticipated. With this in mind, we consider a pessimistic point of view and study a setting where the transportation agencies seek a toll vector that minimizes the worst-case risk.

3.1.1 An Illustrative Example

This section describes the difference between the optimistic and the pessimistic approach in the hazmat toll pricing problem via a simple illustrative example, depicted in Figure 3.1a. There are two paths available: path 1 is $O \to A \to D$, and path 2 is $O \to B \to D$. The shortest path from O to D is path 2, with a cost of 2 and a risk of 20. On the other hand, path 1 is more costly, with a cost of 2, but is safer, with a risk of 10. If the transportation agency wishes to incentivize users to select path 2 while minimizing the financial impact, they may select an optimal toll policy of $t_{BD} = 1$ and $t_{ij} = 0$ for all other links (i, j).

Under this toll policy, both path 1 and path 2 cost the same to network users, i.e., there are two optimal solutions to the inner problem. If we make an *optimistic* assumption about user behavior, we assume that all network users will select path 1, the least risky option. Alternatively, if we make a *pessimistic* assumption, we consider the worst case, where each user selects path 2. Observe that, under this pessimistic approach, increasing the toll t_{BD} slightly, will induce users to always select path 1, as it removes the multiple optimal solutions. However, the upper-level problem will attempt to set the toll $t_{BD} = 1+\epsilon$ for very small $\epsilon > 0$ to lower the financial cost borne by the network users, and so no optimal solution in fact exists for this particular example.

Alternatively, we could work under the assumption that network users are not perfectly rational and are willing to accept solutions that are very slightly suboptimal. Consider that a new path is added as in Figure 3.1b, which we call path 3: directly from O to D. After tolling $t_{BD} = 1$, the cost of paths 1 and 2 is 3, and the cost of path 3 is 3.1, which is very close to 3. While path 3 is not optimal, but it may be sufficiently close to optimal that network users deem it acceptable. Under this behavioral assumption, the optimistic approach will lead us to set a toll of $t_{BD} = 1$, under which policy the network users will choose path 3 with a risk of 0.

Now, suppose that the risk along path 3 is much higher, with a value of 50. If we are optimistic, we believe that the network users can be induced to choose path 1 with risk 10 by tolling $t_{BD} = 1$. If we are pessimistic, however, we believe that the same toll policy will lead to users selecting path 3, with a substantially higher risk of 50. Therefore, to be robust, it would be reasonable to toll $t_{BD} = 2$ and $t_{OD} = 2$, for example, to make sure that path 1 becomes the only acceptable path to the network users.

As we have observed in this example, the optimal toll policy largely depends on the behavioral assumptions. Later, we discuss the existence of optimal solutions under optimistic and pessimistic formulations. We also present an approximation of the pessimistic problem, for which the solution existence is assured.

3.1.2 Related Works

Toll policy to regulate the use of roads for shipping hazardous materials has been proposed by Marcotte et al. (2009). Other studies extend the model proposed by Marcotte et al. (2009) by developing different toll pricing frameworks which control both regular and hazmat traffic (Wang et al., 2012), incorporate the spreading of the risk in the network (Bianco et al., 2015), or permit a nonlinear travel delay function (Esfandeh et al., 2016). The hazmat toll pricing problem is an example of a bilevel optimization problem.

In the upper-level optimization problem, the leader (transportation agency) aims to minimize the risk by tolling a set of roads. In the inner problem, given this set of road tolls, the follower (shipping carrier) chooses the shortest path. Optimizing the upper-level problem depends on the solution of the inner problem which may not be unique. To address this issue, two different approaches have been considered, namely the optimistic and the pessimistic approach. In the optimistic approach, the leader assumes that the follower chooses the best response, while in the pessimistic approach, it is assumed that the follower chooses the worst response. While the previous studies in hazmat toll pricing literature consider an optimistic approach, in this study, we contribute to hazmat toll pricing literature by considering a pessimistic approach. Previous studies have explored the necessary optimality conditions for general pessimistic bilevel programming (Dempe et al., 2014) and developed exact (Lozano and Smith, 2017) and approximate (Wiesemann et al., 2013) solution approaches for solving the pessimistic bilevel problem. In this research, we study the conditions under which optimal solutions exist for both the pessimistic and the *approximate* pessimistic hazmat toll pricing problem. We also develop solution approaches to solve the approximate pessimistic hazmat toll pricing problem, using ideas from Wiesemann et al. (2013) and disjunctive programming.

The underlying assumption in the approximate pessimistic hazmat toll pricing problem is related to the notion of *bounded rationality* and *satisficing behavior* of the network users. When network users are *satisficing*—a term coined by Simon (1956)—they are indifferent among the alternatives whose costs are within a certain threshold. This notion has been actively applied in transportation research (Guo and Liu, 2011; Di and Liu, 2016; Sun et al., 2017; Eikenbroek et al., 2018; Takalloo and Kwon, 2019). Under this setting, any path within a threshold of the shortest path is acceptable to the users and can be the output of the lower level problem. In this study, we assume that the transportation agency seeks a set of link tolls that minimizes the worst-case risk of the acceptable, or satisficing, paths. In other words, the transportation agency assumes that the network users choose the riskiest path among the set of satisficing paths and aims to minimize the transportation hazmat risk in the network by tolling certain roads.

3.1.3 Contributions

To the best of our knowledge, this study considers the pessimistic approach in hazmat toll pricing for the first time. Two main challenges arise regarding the existence and computation of optimal solutions, which are addressed here. First, we show that the approximate pessimistic hazmat toll pricing problem always admits an optimal solution, while the pessimistic problem does not. To prove the existence of optimal solutions, we consider an alternative problem with bounded tolls, which we further leverage to show that the approximate pessimistic problem admits an optimal solution. Second, we present two formulations for the master problem, using techniques from disjunctive programming. We analyze these formulations and show a connection between the "big-M constants" that appear in them with the bounded toll problem. Additionally, we provide valid inequalities for these formulations to improve the computational efficacy of our approach. Finally, we present numerical experiments, where we observe on realistic hazmat networks that a) the optimistic approach can lead to a substantial underestimate of true risk, and that b) our methods can produce high-quality solutions in a reasonable amount of computation time.

3.1.4 Notation

Throughout this chapter, we try to keep consistency in the mathematical notation. Calligraphic upper-case alphabets are used to denote sets of indices such as $\mathcal{S}$ or $\mathcal{A}$. Lower-case alphabets are for indices, as in $s \in \mathcal{S}$, or scalar quantities, as in z or n_s. Lower-case bold alphabets are for vector quantities as in $\mathbf{t}$ or $\mathbf{x}_s$, while their elements are t_{ij} or x_{ijs}. Upper-case bold alphabets are for sets of vectors as in $\mathbf{X}$ or $\mathbf{Q}_s$. Blackboard bold alphabets are for sets of solutions to some optimization problems as in $\mathbb{X}$. We also distinguish scalar zero 0 from vector zero $\mathbf{0}$ in a similar fashion.

3.2 Bilevel Hazmat Toll Pricing Problem

Consider a network $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ with $\mathcal{N}$ as the set of nodes, and $\mathcal{A}$ as the set of links. We denote the set of origin-destination shipments in the network by $\mathcal{S}$. Each origin-destination shipment $s \in \mathcal{S}$ has a fixed hazmat type $h(s) \in \mathcal{H}$. Moreover, $n_s \geq 1$ denote the number of trucks needed for shipment $s \in \mathcal{S}$. Parameter $c_{ij} \geq 0$ denote the length or the cost of link $(i,j) \in \mathcal{A}$ and $\rho_{ij}^{h(s)} \geq 0$ represents the risk exposure on link (i,j) when hazmat type h is carried. Constant b_{si} is defined as 1 if $i = o(s)$, -1 if $i = d(s)$, or 0 otherwise for all $i \in \mathcal{N}, s \in \mathcal{S}$. With this notation, we can present the (optimistic) bilevel Hazmat Toll Pricing (HTP) problem as follows:

$$
\begin{aligned}
\text{HTP}: \quad & \underset{\mathbf{t}}{\text{minimize}} && \sum_{s\in\mathcal{S}} \sum_{(i,j)\in\mathcal{A}} n_s \Big(\rho_{ij}^{h(s)} + \alpha\Big(c_{ij} + t_{ij}^{h(s)}\Big)\Big) x_{sij} \\
& \text{subject to} && t_{ij}^{h(s)} \geq 0 && \forall (i,j) \in \mathcal{A}, s \in \mathcal{S} \\
& \text{where, for each } s \in \mathcal{S},\ \mathbf{x}_s \text{ solves:} \\
& \underset{\mathbf{x}_s}{\text{minimize}} && \sum_{(i,j)\in\mathcal{A}} \Big(c_{ij} + t_{ij}^{h(s)} + \beta\rho_{ij}^{h(s)}\Big) x_{sij} \\
& \text{subject to} && \sum_{(i,j)\in\mathcal{A}} x_{sij} - \sum_{(j,i)\in\mathcal{A}} x_{sji} = b_{si} && \forall i \in \mathcal{N}
\end{aligned}
$$

$$x_{sij} \in \{0, 1\} \qquad \forall (i,j) \in \mathcal{A},$$

where α and β are nonnegative constants. The upper-level decision variable t_{ij}^h is the toll imposed on link (i,j) for hazmat type h, and the lower level decision variable x_{sij} is 1 if link (i,j) is used for carrying shipment s and is 0 otherwise. Note that the above problem is separable for each hazmat type $h \in \mathcal{H}$ since we charge the same toll $t_{ij}^{h(s)}$ if the hazmat type is the same. Once we define $\mathcal{S}_h = \{s \in \mathcal{S} : h(s) = h\}$ as the set of shipments of each hazmat type $h \in \mathcal{H}$, we can write the problem for each h and $\mathcal{S}_h$. Therefore, without the loss of generality, we can drop superscripts h and $h(s)$ from the problem for the simplicity of the presentation.

3.3 Optimistic and Pessimistic Formulations

Let us define the solution set of inner problem for each shipment $s \in \mathcal{S}$ separately as:

$$\mathbf{X}_s = \left\{ \mathbf{x}_s \;\middle|\; \sum_{(i,j)\in\mathcal{A}} x_{sij} - \sum_{(j,i)\in\mathcal{A}} x_{sji} = b_{si} \;\forall i \in \mathcal{N}, \quad x_{sij} \in \{0,1\} \;\forall (i,j) \in \mathcal{A} \right\},$$

where each vector $\mathbf{x}_s \in \mathbf{X}_s$ denotes a path for shipment $s \in \mathcal{S}$. Note that this set is independent of toll vector $\mathbf{t} = (t_{ij} : (i,j) \in \mathcal{A})$. In overloading of notation, we will also use $\mathbf{x}_s$ to denote a sub-vector of $\mathbf{x}$ associated with shipment $s \in \mathcal{S}$. Therefore, we can write $\mathbf{X} = \{\mathbf{x} \mid \mathbf{x}_s \in \mathbf{X}_s \;\forall s \in \mathcal{S}\}$. As the lower level problem for a given $\mathbf{t}$ is a shortest path problem, when we consider $\mathbf{X}$ in the inner problem of hazmat carriers, subtour elimination constraints are unnecessary and we can relax the integrality conditions to interval conditions $[0, 1]$.

For simplicity of notation, we define a shorthand for the lower-level objective function:

$$h_s(\mathbf{t}, \mathbf{x}_s) := \sum_{(i,j)\in\mathcal{A}} (c_{ij} + t_{ij} + \beta\rho_{ij}) x_{sij}$$

for each $s \in \mathcal{S}$. The set of optimal solutions to the inner problem for a given toll policy is given by the following point-to-set mappings:

$$\mathbb{X}_s(\mathbf{t}) := \underset{\mathbf{x}_s \in \mathbf{X}_s}{\arg\min}\ h_s(\mathbf{t}, \mathbf{x}_s) \qquad \forall s \in \mathcal{S}$$
$$\mathbb{X}(\mathbf{t}) := \underset{\mathbf{x} \in \mathbf{X}}{\arg\min} \sum_{s \in \mathcal{S}} h_s(\mathbf{t}, \mathbf{x}_s) = \prod_{s \in \mathcal{S}} \mathbb{X}_s(\mathbf{t}).$$

Similarly, for the upper-level objective function we define

$$f_s(\mathbf{t}, \mathbf{x}_s) := \sum_{(i,j) \in \mathcal{A}} n_s \Big(\rho_{ij} + \alpha (c_{ij} + t_{ij}) \Big) x_{sij}$$
$$f(\mathbf{t}, \mathbf{x}) := \sum_{s \in \mathcal{S}} f_s(\mathbf{t}, \mathbf{x}_s).$$

Then, we can rewrite HTP compactly as follows:

$$\text{HTP}: \quad \underset{\mathbf{t} \geq \mathbf{0},\ \mathbf{x} \in \mathbb{X}(\mathbf{t})}{\text{minimize}} \quad f(\mathbf{t}, \mathbf{x})$$

The above HTP problem is called an optimistic formulation in the bilevel optimization literature. When $\alpha = 0$, we can solve this problem by a two-phase method based on inverse optimization (Marcotte et al., 2009). The inverse problem always has a feasible solution, and the optimal toll solution exists.

When $\alpha \neq 0$, HTP becomes a bilevel optimization problem. Marcotte et al. (2009) propose a single level optimization formulation for bilevel optimization by substituting the lower level problem with KKT condition. Since the lower level problem can have multiple optimal solutions, it is unclear which path will be chosen by the followers. This optimistic formulation assumes that the hazmat carriers will choose the least risky path(s) as discussed for the simple example in Figure 3.1a. Objective function of HTP involves a bilinear term $t_{ij} x_{sij}$. By linearizing these bilinear terms, HTP can be cast as a bilevel mixed-integer linear program.

In the pessimistic approach, it is assumed that hazmat carriers are indifferent among multiple optimal solutions and choose the riskiest path. We write the pessimistic formulation as follows:

$$\text{PHTP}: \quad \underset{\mathbf{t}\geq\mathbf{0}}{\text{minimize}} \quad \max_{\mathbf{x}\in\mathbb{X}(\mathbf{t})} \quad f(\mathbf{t},\mathbf{x}). \tag{3.1a}$$

By introducing auxiliary variables z, we can also write equivalently:

$$\text{PHTP}: \quad \underset{\mathbf{t}\geq\mathbf{0},\ z\geq 0}{\text{minimize}} \quad z \tag{3.2a}$$

$$\text{subject to} \quad f(\mathbf{t},\mathbf{x}) - z \leq 0 \qquad \forall \mathbf{x} \in \mathbb{X}(\mathbf{t}), \tag{3.2b}$$

or equivalently, by letting $z = \sum_{s\in\mathcal{S}} z_s$ and $\mathbf{z} = (z_s : s \in \mathcal{S})$:

$$\text{PHTP}: \quad \underset{\mathbf{t}\geq\mathbf{0},\ \mathbf{z}\geq\mathbf{0}}{\text{minimize}} \quad \sum_{s\in\mathcal{S}} z_s$$

$$\text{subject to} \quad f_s(\mathbf{t},\mathbf{x}_s) - z_s \leq 0 \qquad \forall \mathbf{x}_s \in \mathbb{X}_s(\mathbf{t}), s \in \mathcal{S}.$$

By definition of $f_(\cdot,\cdot,)$, both z and $\mathbf{z}$ are nonnegative. Furthermore, we consider an approximation to PHTP by relaxing the optimizing-behavior assumption of the hazmat carriers (Wiesemann et al., 2013). Define

$$\mathbb{X}_s^{\epsilon}(\mathbf{t}) := \big\{\mathbf{y}_s \in \mathbf{X}_s | h_s(\mathbf{t},\mathbf{y}_s) < h_s(\mathbf{t},\mathbf{y}'_s) + \epsilon_s \quad \forall \mathbf{y}'_s \in \mathbf{X}_s\big\}, \tag{3.3}$$

$$\mathbb{X}^{\epsilon}(\mathbf{t}) := \big\{\mathbf{y} \in \mathbf{X} | \mathbf{y}_s \in \mathbb{X}_s^{\epsilon}(\mathbf{t}) \quad \forall s \in \mathcal{S}\big\}, \tag{3.4}$$

where $\mathbf{y}_s$ and $\mathbf{y}'_s$ denote a sub-vector of $\mathbf{y}$ and $\mathbf{y}'$, respectively, for shipment $s \in \mathcal{S}$. Call any lower-level solution $\mathbf{y} \in \mathbb{X}^{\epsilon}(\mathbf{t})$ an ϵ-*approximate* solution under $\mathbf{t}$. Similarly, $\mathbf{y}_s$ represents an ϵ-*approximate* path for shipment $s \in \mathcal{S}$. The ϵ-approximation to PHTP is defined as follows:

$$\text{PHTP}_\epsilon : \quad \underset{\mathbf{t}\geq\mathbf{0},\ \mathbf{z}\geq\mathbf{0}}{\text{minimize}} \quad \sum_{s\in\mathcal{S}} z_s \tag{3.5a}$$

$$\text{subject to} \quad f_s(\mathbf{t}, \mathbf{x}_s) - z_s \leq 0 \qquad \forall \mathbf{x}_s \in \mathbb{X}_s^\epsilon(\mathbf{t}), s \in \mathcal{S} \tag{3.5b}$$

The ϵ-approximation notion is related to the satisficing behavior and the bounded rationality of network users (Sun et al., 2017). The satisficing network users may choose any path whose cost is close enough to the shortest path. which is related to (3.3). In this study, we use the satisficing path and ϵ-*approximate* path interchangeably.

3.4 Existence of Optimal Toll Solutions

In this section, we discuss the existence of an optimal toll solution for the optimistic, pessimistic, and $\boldsymbol{\epsilon}$-approximate pessimistic formulation.

Let us consider the example in Figure 3.1a. We assume that $\beta = 0$ for simplicity and $\alpha > 0$ is sufficiently small so that the risk component is dominant in the upper-level objective function. We consider shipment 1 from O to D; shipment 2 from O to B; and shipment 3 from B to D. We assume $n_s = 1$ for each shipment $s \in \{1, 2, 3\}$. Although there are no path-change opportunities for shipments 2 and 3, they still contribute to the objective function since $\alpha > 0$. Clearly there are two paths for shipment 1:

$$\text{path 1} : O \to A \to D$$

$$\text{path 2} : O \to B \to D$$

The upper-level objective function for path 1 and path 2 is as follows:

$$\text{path 1} : [5 + \alpha(1 + t_{OA})] + [5 + \alpha(2 + t_{AD})]$$

$$\text{path 2} : [10 + \alpha(1 + t_{OB})] + [10 + \alpha(1 + t_{BD})].$$

the lower level objective function becomes

$$\text{path } 1 : [(1 + t_{OA}) + (2 + t_{AD})] = 3 + t_{OA} + t_{AD}$$

$$\text{path } 2 : [(1 + t_{OB}) + (1 + t_{BD})] = 2 + t_{OB} + t_{BD}$$

For the optimistic formulation HTP, any toll vector such that $3+t_{OA}+t_{AD} \leq 2+t_{OB}+t_{BD}$ is sufficient to make path 1 optimal for the lower-level problem. Since $\alpha > 0$, we require $3 + t_{OA} + t_{AD} = 2 + t_{OB} + t_{BD}$; for example

$$t^*_{BD} = 1,\ t^*_{OA} = t^*_{AD} = t^*_{OB} = 0.$$

is a solution for HTP.

For the pessimistic formulation PHTP, however, we require $3+t_{OA}+t_{AD} < 2+t_{OB}+t_{BD}$ to make path 1 the *unique* optimal solution to the inner problem. We need to solve the following problem:

$$\begin{aligned} \underset{\mathbf{t}}{\text{minimize}} \quad & [5 + \alpha(1 + t_{OA})] + [5 + \alpha(2 + t_{AD})] \\ \text{subject to} \quad & 3 + t_{OA} + t_{AD} < 2 + t_{OB} + t_{BD} \\ & \mathbf{t} \geq \mathbf{0}. \end{aligned}$$

Since the constraint is a strict inequality and α is positive, there exists no optimal solution.

We can, however, show that PHTP$_\epsilon$ has an optimal solution for the same example. Suppose $\epsilon_s = 0.1$. To make path 1 the *unique* ϵ-approximate solution to the inner problem, we require both

$$3 + t_{OA} + t_{AD} < 2 + t_{OB} + t_{BD} + 0.1,$$

$$2 + t_{OB} + t_{BD} \nless 3 + t_{OA} + t_{AD} + 0.1.$$

We need to solve the following problem:

$$\begin{aligned}
\underset{\mathbf{t}}{\text{minimize}} \quad & [5+\alpha(1+t_{OA})]+[5+\alpha(2+t_{AD})] \\
\text{subject to} \quad & 3+t_{OA}+t_{AD} < 2+t_{OB}+t_{BD}+0.1, \\
& 2+t_{OB}+t_{BD} \geq 3+t_{OA}+t_{AD}+0.1. \\
& \mathbf{t} \geq \mathbf{0}.
\end{aligned}$$

Thus, an optimal toll policy under PHTP_ϵ is

$$t_{BD}=1.1,\ t_{OA}=t_{AD}=t_{OB}=0.$$

for any $\alpha \geq 0$. The existence of an optimal solution is guaranteed for this example, mainly because of the strict inequality in the definition of $\mathbb{X}(\cdot)$ in (3.4).

As shown in the example above, the pessimistic formulation PHTP may not have an optimal solution, although there always exists an optimal solution to the optimistic formulation HTP.

Proposition 3.1. The optimistic formulation HTP admits an optimal solution for any $\alpha \geq 0$. However, there exists no optimal solution to PHTP, in general.

Unlike the pessimistic formulation PHTP, the ϵ-approximate pessimistic problem PHTP_ϵ guarantees a solution. To prove the optimal solution existence for PHTP_ϵ, first, we consider the case when the toll vector $\mathbf{t}$ is bounded from above. When $\mathbf{t}$ is bounded, z is also bounded. In particular, we assume that t_{ij} is bounded as follows:

$$t_{ij} \leq B_{ij} := \frac{z_{\mathbf{t}=\mathbf{0}}}{\alpha} + \max_{s\in\mathcal{S}} \epsilon_s + \beta \sum_{(i,j)\in\mathcal{A}} \rho_{ij} \qquad \forall (i,j)\in\mathcal{A}, \tag{3.6}$$

where $z_{\mathbf{t}=\mathbf{0}}$ is an upper bound for PHTP_ϵ obtained by setting $\mathbf{t}=\mathbf{0}$. We denote this bounded version of PHTP_ϵ by $\overline{\text{PHTP}_\epsilon}$ and prove that there is an optimal solution for this problem.

Lemma 3.1. There exists an optimal solution to $\overline{\text{PHTP}_\epsilon}$ for all $\epsilon > \mathbf{0}$.

In fact, Lemma 3.1 holds not only for B_{ij} but also for any finite upper bound for t_{ij}. Therefore, when t_{ij} has an exogenous upper bound, as in the second-best toll pricing (Lawphongpanich and Hearn, 2004), Lemma 3.1 assures the existence of ϵ-approximate pessimistic toll pricing for all $\epsilon \geq \mathbf{0}$ and $\alpha \geq 0$. When there is no exogenous upper bound on t_{ij}, we rely on the endogenous bound, B_{ij}, given in (3.6), which requires $\alpha > 0$. Using the endogenous bound B_{ij} and Lemma 3.1, we show that for any given solution to the original problem PHTP_ϵ, there exists a solution to the bounded problem $\overline{\text{PHTP}_\epsilon}$ that is no worse than the given solution, which leads to the existence of ϵ-approximate pessimistic toll.

Proposition 3.2. There exists an optimal solution to PHTP_ϵ for all $\epsilon > \mathbf{0}$ and $\alpha > 0$.

As discussed above, the pessimistic formulation PHTP may not have an optimal solution. For this reason, we consider the approximate pessimistic formulation PHTP_ϵ with $\boldsymbol{\epsilon} > \mathbf{0}$ and develop solution approaches for it in the upcoming sections.

3.5 Solution Approaches for the Approximate Formulation

In this section, we reformulate PHTP_ϵ as a semi-infinite program and devise a cutting plane algorithm based on the method of Wiesemann et al. (2013).

First, note that for a given $s \in \mathcal{S}$, constraint (3.5b) is equivalent to the following disjunctive constraint (Wiesemann et al., 2013, Proposition 4.1):

$$\begin{aligned} &\left[\mathbf{x} \in \mathbb{X}_s^\epsilon(\mathbf{t}) \implies f_s(\mathbf{t}, \mathbf{x}_s) - z \leq 0\right] && \forall \mathbf{x} \in \mathbf{X} \\ \iff &\left[\mathbf{x} \notin \mathbb{X}_s^\epsilon(\mathbf{t})\right] \quad \vee \quad \left[f_s(\mathbf{t}, \mathbf{x}_s) - z \leq 0\right] && \forall \mathbf{x} \in \mathbf{X}, \end{aligned}$$

which is also equivalent to the following statement:

$$\exists \mathbf{y} \in \mathbf{X} : \quad [h_s(\mathbf{t}, \mathbf{x}_s) \geq h_s(\mathbf{t}, \mathbf{y}_s) + \epsilon_s] \quad \vee \quad [f_s(\mathbf{t}, \mathbf{x}_s) - z_s \leq 0] \qquad \mathbf{x} \in \mathbf{X}$$

Hence, PHTP_ϵ is equivalent to the following problem with disjunctive constraints:

$$\underset{\mathbf{t}\geq\mathbf{0},\ \mathbf{z}\geq\mathbf{0},\ \mathbf{y}\in\mathbf{X}}{\text{minimize}} \quad \sum_{s\in\mathcal{S}} z_s \tag{3.7a}$$

$$\text{subject to} \quad [h_s(\mathbf{t},\mathbf{x}_s) \geq h_s(\mathbf{t},\mathbf{y}_s) + \epsilon_s] \quad \vee \quad [f_s(\mathbf{t},\mathbf{x}_s) - z_s \leq 0] \quad \forall s\in\mathcal{S}, \mathbf{x}\in\mathbf{X} \tag{3.7b}$$

If we assume that there is at least one path for each shipment, then the lower level problem has a solution and $\mathbb{X}^\epsilon(\mathbf{t}) \neq \emptyset$. Under this assumption, variable z_s attains the worst-case objective value $f_s(.)$ of the lower level problem for the shipment s at optimality. Next, we present a cutting plane algorithm for solving (3.7).

3.5.1 Cutting Plane Algorithm For Approximate Pessimistic Problem

In this section, we describe an iterative algorithm for solving PHTP_ϵ, which is based on the iterative algorithm proposed by Wiesemann et al. (2013). In each iteration k of the algorithm, we solve two problems, namely the master problem Master_k and the subproblem Sub_k. Master_k is a relaxation of (3.7), and considers only a finite set of paths, namely $\mathbf{X}^k \subset \mathbf{X}$, and can be formulated as follows:

Master_k :

$$\underset{\mathbf{t}\geq\mathbf{0},\ \mathbf{z}\geq\mathbf{0},\ \mathbf{y}\in\mathbf{X}}{\text{minimize}} \quad \sum_{s\in\mathcal{S}} z_s \tag{3.8a}$$

$$\text{subject to} \quad [h_s(\mathbf{t},\mathbf{x}^l_s) \geq h_s(\mathbf{t},\mathbf{y}_s) + \epsilon_s] \quad \vee \quad [f_s(\mathbf{t},\mathbf{x}^l_s) - z_s \leq 0] \quad \forall s\in\mathcal{S}, \mathbf{x}^l\in\mathbf{X}^k, \tag{3.8b}$$

which produces an optimal solution $(\mathbf{t}^k, \mathbf{z}^k, \mathbf{y}^k)$. We formulate Master_k as a MILP in Section 3.5.3 and 3.5.4. Next, we can formulate the subproblem Sub_k as follows:

$$\text{Sub}_k : \underset{\mathbf{x}\in\mathbf{X},\ \mathbf{w}}{\text{maximize}} \quad \sum_{s\in\mathcal{S}} w_s$$

$$\text{subject to} \quad w_s \leq h_s^k - h_s(\mathbf{t}^k, \mathbf{x}_s) + \epsilon_s \qquad \forall s \in \mathcal{S}$$

$$w_s \leq f_s(\mathbf{t}^k, \mathbf{x}_s) - z_s^k \qquad \forall s \in \mathcal{S}$$

where h_s^k is the minimum lower-level objective function value under $\mathbf{t}^k$, given by

$$h_s^k = \min_{\mathbf{y} \in \mathbf{X}} h_s(\mathbf{t}^k, \mathbf{y}_s). \tag{3.9}$$

To solve Sub_k, we need both subtour elimination constraints and integrality conditions in $\mathbf{X}$. Note that subproblem Sub_k finds a path solution $\mathbf{x}_s^k \in \mathbb{X}_s(\mathbf{t}^k)$ for each shipment s in $\mathcal{S}$. We use this solution to generate some cuts, which are added to Master_k. For the optimal solution $\mathbf{w}^k$ to Sub_k, suppose there exists a shipment such that $w_s^k > 0$. Then the following inequalities hold:

$$h_s(\mathbf{t}^k, \mathbf{x}_s^k) < h_s^k + \epsilon_s,$$

$$z_s^k < f_s(\mathbf{t}^k, \mathbf{x}_s^k),$$

which imply that there exists an ϵ-approximate path with a higher upper-level objective function under toll vector $\mathbf{t}^k$ than z_s^k. In this case, we update the master problem by $\mathbf{X}^{k+1} = \mathbf{X}^k \cup \{\mathbf{x}^k\}$. If $\mathbf{w}^k \leq \mathbf{0}$, then there is no ϵ-approximate solution $\mathbf{x}^k$ with a higher risk than current set of path solutions $\mathbf{X}^k$ and $\mathbf{t}^k$ is optimal. A summary of the algorithm is provided.

Note that the proposed cutting plane algorithm is an exact approach. The number of constraints in Master_k grows as the algorithm iterates. Therefore, Master_k becomes difficult to solve. As a result, it may not be possible to solve the approximate pessimistic formulation PHTP_ϵ optimally for large networks. However, we can obtain a primal bound using the toll vector $\mathbf{t}^k$ that we obtained by solving Master_k. Since ϵ_s is small relative to the value of $h_s(\mathbf{t}, \mathbf{x}_s)$, then the number of paths within $\mathbb{X}_\epsilon(\mathbf{t}^k)$ is small. Therefore we can use Yen's K-shortest path algorithm (Yen, 1971) to enumerate the set of approximately optimal paths.

Algorithm 3: Cutting Plane Algorithm For PHTP_ϵ

Input: Network $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ and parameters $\mathbf{c}, \mathbf{n}, \boldsymbol{\rho}, \alpha, \beta$
Output: Optimal toll $\mathbf{t}^*$
Initialization. Set iteration counter $k = 0$ and $\mathbf{X}_0 = \emptyset$;
while $\mathbf{w}^k > \mathbf{0}$ **do**
  Given $\mathbf{X}^k$, solve Master_k to obtain optimal solution $(\mathbf{t}^k, \mathbf{z}^k, \mathbf{y}^k)$;
  Given $\mathbf{t}^k$, compute h_s^k by solving (3.9) ;
  Given $\mathbf{t}^k$ and $\mathbf{z}^k$, solve subproblem Sub_k to obtain optimal solution $(\mathbf{x}^k, \mathbf{w}^k)$;
  Set $\mathbf{X}^{k+1} = \mathbf{X}^k \cup \{\mathbf{x}^k\}$ and $k \leftarrow k + 1$.
end

Once we enumerate those paths, we can compute the upper-level objective function $f_s(\mathbf{t}^k, \mathbf{x}_s)$ for them and choose the one with maximum $f_s(\mathbf{t}^k, \mathbf{x}_s)$, which gives us a primal bound for PHTP_ϵ. In each iteration of the algorithm, we obtain a dual bound for the problem by solving Master_k, which can be used to calculate the optimality gap for the primal bound.

In each iteration of the algorithm, instead of adding only one path to the Master_k for each shipment s, we can add multiple paths. In particular, to generate multiple paths in iteration k of the algorithm, we find the set of ϵ-approximate shortest paths under $\mathbf{t}_k$, compute $f_s(\cdot)$ for them, and add those with $f_s(\cdot)$ higher than z_s^k to the problem.

3.5.2 Analysis of Master_k

As the following proposition shows, we can simply ignore subtour elimination constraints in $\mathbf{X}$ in Master_k.

Proposition 3.3. For any solution $\mathbf{y}^k \in \mathbf{X}$ with subtours for Master_k, there exists a solution $\overline{\mathbf{y}} \in \mathbf{X}$ without subtours, at least as good as $\widehat{\mathbf{y}}$.

To solve PHTP_ϵ efficiently, we need to be able to solve Master_k in a reasonable amount of time. In this section, we provide some valid cuts for PHTP_ϵ, which can be added to Master_k. Adding these cuts to Master_k can potentially decrease the solution time of Master_k. We consider the following optimization problem to provide valid cuts:

$$\underset{\mathbf{t}\geq\mathbf{0},\ \mathbf{x}\in\mathbf{X}}{\text{minimize}} \quad \sum_{s\in\mathcal{S}} f_s(\mathbf{t},\mathbf{x}_s) = \sum_{s\in\mathcal{S}}\sum_{(i,j)\in\mathcal{A}} n_s\Big(\rho_{ij}+\alpha(c_{ij}+t_{ij})\Big)x_{sij} \tag{3.10}$$

Since $\mathbf{X}$ is independent from $\mathbf{t}$ and Problem (3.10) is a minimization problem, and therefore, we will necessarily have $\mathbf{t}=\mathbf{0}$ at optimality. Hence, we can rewrite (3.10) as:

$$R^{\mathbf{0}} := \min_{\mathbf{x}\in\mathbf{X}} \sum_{s\in\mathcal{S}} f_s(\mathbf{0},\mathbf{x}_s) = \min_{\mathbf{x}\in\mathbf{X}} \sum_{s\in\mathcal{S}}\sum_{(i,j)\in\mathcal{A}} n_s\Big(\rho_{ij}+\alpha c_{ij}\Big)x_{sij}$$

Similarly, we can define $R^{\mathbf{0}}_s$ for each shipment s as:

$$R^{\mathbf{0}}_s := \min_{\mathbf{x}\in\mathbf{X}} \sum_{s\in\mathcal{S}}\sum_{(i,j)\in\mathcal{A}} n_s\Big(\rho_{ij}+\alpha c_{ij}\Big)x_{sij}$$

Then, we have the following valid inequality for PHTP_ϵ:

$$z_s \geq R^{\mathbf{0}}_s \qquad \forall s\in\mathcal{S}.$$

Consequently we have $z=\sum_{s\in\mathcal{S}} z_s \geq R^{\mathbf{0}}$. Moreover, as described in the previous section, by setting the toll vector to any nonnegative vector $\bar{\mathbf{t}}$ and using Yen's K-shortest path algorithm, we can obtain an upper bound for PHTP_ϵ, which we denote by $U_{\bar{\mathbf{t}}}$. Therefore, we have

$$z=\sum_{s\in\mathcal{S}} z_s \leq U_{\bar{\mathbf{t}}} \tag{3.11}$$

We can add this valid cuts to Master_k to obtain a tighter and a stronger relaxation.

Linearizing the bilinear term $t_{ij}y_{sij}$ in (3.8b), we can reformulate Master_k as follows:

$$\underset{\mathbf{t},\mathbf{z},\mathbf{y},\boldsymbol{\tau}}{\text{minimize}} \sum_{s\in\mathcal{S}} z_s \tag{3.12a}$$

$$\left[\sum_{(i,j)\in\mathcal{A}}(c_{ij}+\beta\rho_{ij}+t_{ij})x^l_{sij}\geq\sum_{(i,j)\in\mathcal{A}}(c_{ij}+\beta\rho_{ij})y_{sij}+\tau_{sij}+\epsilon_s\right]\vee\left[f_s(\mathbf{t},\mathbf{x})-z_s\leq 0\right]$$

$$\forall s\in\mathcal{S}, l\in[\![|\mathbf{X}^k|]\!] \qquad (3.12b)$$

$$\tau_{sij}\leq M^3_{sij}y_{sij} \qquad \forall s\in\mathcal{S},(i,j)\in\mathcal{A} \qquad (3.12c)$$

$$\tau_{sij}\leq t_{ij} \qquad \forall s\in\mathcal{S},(i,j)\in\mathcal{A} \qquad (3.12d)$$

$$\tau_{sij}\geq t_{ij}+M^3_{sij}(y_{sij}-1) \qquad \forall s\in\mathcal{S},(i,j)\in\mathcal{A} \qquad (3.12e)$$

$$\mathbf{t}\geq\mathbf{0},\mathbf{z}\geq\mathbf{0},\boldsymbol{\tau}\geq\mathbf{0} \qquad (3.12f)$$

$$\mathbf{y}\in\mathbf{X} \qquad (3.12g)$$

where $[\![|\mathbf{X}^k|]\!]=\{1,2,...,|\mathbf{X}^k|\}$. Constraints (3.12c)–(3.12e) linearize $t_{ij}y_{ij}$. Marcotte et al. (2009) propose a big-M formulation for optimistic problem. They use the longest path for bounding the maximum toll that can be charged on an arc. The length of the longest path, unfortunately, cannot be used for bounding the toll on each link. We present several lemmas and propositions to provide valid bounds on M^3_{sij} in (3.12c) and (3.12e). In these lemmas and propositions, we assume that there should always be at least one ϵ-approximation solution for each shipment $s\in\mathcal{S}$ in Master$_k$. In other words, there always exists $l(s)\in[\![|\mathbf{X}^k|]\!]$ for each $s\in\mathcal{S}$ such that $z^k_s=f_s(\mathbf{t}^k,\mathbf{x}^{l(s)}_s)$. First, based on this assumption, we provide a bound on the amount of toll charged on shipment $s\in\mathcal{S}$ under an optimal solution to Master$_k$. We denote the amount of toll charged by T^k_s.

Lemma 3.2. Let $(\mathbf{t}^k,\mathbf{z}^k,\mathbf{y}^k)$ be an optimal solution to **Master**$_k$. Suppose there exists $l(s)\in[\![|\mathbf{X}^k|]\!]$ for each $s\in\mathcal{S}$ such that $z^k_s=f_s(\mathbf{t}^k,\mathbf{x}^{l(s)}_s)$. Let $T^k_s=\sum_{(i,j)\in\mathcal{A}}t^k_{ij}x^{l(s)}_{sij}$. Then, the following holds:

$$T^k_s\leq\overline{T}_s:=\frac{U_{\bar{\mathbf{t}}}-R^{\mathbf{0}}}{\alpha n_s}$$

for any nonnegative toll vector $\bar{\mathbf{t}}\geq\mathbf{0}$.

Lemma 3.3. Let $(\mathbf{t}^k, \mathbf{z}^k, \mathbf{y}^k)$ be an optimal solution to $\mathbf{Master}_k$. Let us define

$$\widehat{\mathbf{y}}(\mathbf{t}^k) := \arg\min_{\mathbf{x}\in\mathbf{X}} \sum_{s\in\mathcal{S}} h_s(\mathbf{t}^k, \mathbf{x}_s) \tag{3.13}$$

Then, $(\mathbf{t}^k, \mathbf{z}^k, \widehat{\mathbf{y}}(\mathbf{t}^k))$ is also an optimal solution to $\mathbf{Master}_k$.

To bound M^3_{sij}, we first define:

$$C^0_s := \max_{\mathbf{x}_s\in\mathbf{X}_s} \quad \sum_{(i,j)\in\mathcal{A}} (c_{ij} + \beta\rho_{ij})x_{sij} \qquad \forall s \in \mathcal{S}, \tag{3.14}$$

where subtour elimination constraints are implicitly assumed in $\mathbf{X}_s$. Essentially, C^0_s is the length of the "longest path" for shipment $s \in \mathcal{S}$ with link cost being $(c_{ij} + \beta\rho_{ij})$ for each $(i, j) \in \mathcal{A}$.

Proposition 3.4. Let $(\mathbf{t}^k, \mathbf{z}^k, \mathbf{y}^k)$ be an optimal solution to $\mathbf{Master}_k$. Suppose there exists $l(s) \in [\![|\mathbf{X}^k|]\!]$ for each $s \in \mathcal{S}$ such that $z^k_s = f_s(\mathbf{t}^k, \mathbf{x}^{l(s)}_s)$. Let $T^k_s = \sum_{(i,j)\in\mathcal{A}} t^k_{ij} x^{l(s)}_{sij}$. Then, the following lower bound for M^{3s}_{ij} is valid in $\mathbf{Master}_k$:

$$M^3_{sij} \geq \max_{s'\in\mathcal{S}} \{C^0_{s'} + \overline{T}_{s'} + \epsilon_{s'}\} - \beta\rho_{ij} - c_{ij} \quad \forall s \in \mathcal{S}. \tag{3.15}$$

Note that we have bounded big-M values in Master_k based on the existence of an ϵ_s-*approximate* path for each shipment. This assumption may put some restriction on the value of $\mathbf{t}^k$ and consequently on the cuts generated by Sub_k in iteration k. Although this assumption generally does not necessarily hold, especially in early iterations, it holds under the optimal solution to PHTP_ϵ in the proposed cutting plane algorithm, when $\mathbf{w}^k \leq \mathbf{0}$.

Lemma 3.4. Suppose on iteration k, we have $\mathbf{w}^k \leq \mathbf{0}$ for $\mathbf{Sub}_k$. Then, there exists at least one ϵ_s-*approximate* path for each shipment $s \in \mathcal{S}$.

3.5.3 Big-M Formulation of the Master Problem

Note that the disjunctive constraints (3.8b) in Master_k can be written with big-M constraints and auxiliary binary variables. We can formulate the master problem as the following MILP:

$$\underset{\mathbf{t},\mathbf{z},\mathbf{y},\mathbf{v},\boldsymbol{\tau}}{\text{minimize}} \quad \sum_{s\in\mathcal{S}} z_s \tag{3.16a}$$

subject to

$$\sum_{(i,j)\in\mathcal{A}} (c_{ij} + \beta\rho_{ij} + t_{ij}) x^l_{sij} \geq \sum_{(i,j)\in\mathcal{A}} \Big[(c_{ij} + \beta\rho_{ij}) y_{sij} + \tau_{sij}\Big] + \epsilon_s - M^{1l}_s v^l_s \quad \forall s \in \mathcal{S}, l \in [\![|\mathbf{X}^k|]\!] \tag{3.16b}$$

$$\sum_{(i,j)\in\mathcal{A}} n_s \Big(\rho_{ij} + \alpha(c_{ij} + t_{ij})\Big) x^l_{sij} - z_s \leq M^{2l}_s (1 - v^l_s) \quad \forall s \in \mathcal{S}, l \in [\![|\mathbf{X}^k|]\!] \tag{3.16c}$$

$$\tau_{sij} \leq M^3_{sij} y_{sij} \quad \forall s \in \mathcal{S}, (i,j) \in \mathcal{A} \tag{3.16d}$$

$$\tau_{sij} \leq t_{ij} \quad \forall s \in \mathcal{S}, (i,j) \in \mathcal{A} \tag{3.16e}$$

$$\tau_{sij} \geq t_{ij} + M^3_{sij}(y_{sij} - 1) \quad \forall s \in \mathcal{S}, (i,j) \in \mathcal{A} \tag{3.16f}$$

$$v^l_s \in \{0,1\} \quad \forall s \in \mathcal{S}, l \in [\![|\mathbf{X}^k|]\!] \tag{3.16g}$$

$$\mathbf{t} \geq \mathbf{0}, \mathbf{z} \geq \mathbf{0}, \boldsymbol{\tau} \geq \mathbf{0} \tag{3.16h}$$

$$\mathbf{y} \in \mathbf{X} \tag{3.16i}$$

Problem Master_k contains big-M constants. The value of such constants has an impact on the solution process. In this section, we seek to find lower bounds for big-M values for Master_k.

Proposition 3.5. Let $(\mathbf{t}^k, \mathbf{z}^k, \mathbf{y}^k, \mathbf{v}^k, \boldsymbol{\tau}^k)$ be an optimal solution to $\mathbf{Master}_k$ in (3.16). Suppose there exists $l(s) \in [\![|\mathbf{X}^k|]\!]$ for each $s \in \mathcal{S}$ such that $z^k_s = f_s(\mathbf{t}^k, \mathbf{x}^{l(s)}_s)$. Let $T^k_s =$

$\sum_{(i,j)\in\mathcal{A}} t_{ij}^k x_{sij}^{l(s)}$. Then the following lower bounds for M_s^{1l} and M_s^{2l} are valid for $\mathbf{Master}_k$:

$$M_s^{1l} \geq C_s^0 + \overline{T}_s + \epsilon_s \tag{3.17}$$

$$M_s^{2l} \geq \sum_{(i,j)\in\mathcal{A}} n_s\Big(\rho_{ij}^s + \alpha(\max_{s'\in\mathcal{S}} \{C_{s'}^0 + \overline{T}_{s'} + \epsilon_{s'}\} - \beta \min_{s'\in\mathcal{S}}\{\rho_{ij}^{s'}\})\Big) \tag{3.18}$$

for all $s \in \mathcal{S}, l \in [\![|\mathbf{X}^k|]\!]$.

3.5.4 Balas Extended Reformulation of the Master Problem

In this subsection we derive another MILP formulation for (3.8b) in Master_k using techniques introduced by Balas (1985). The construction of Balas emits MILP formulations for disjunctive sets that uses additional continuous variables and is *ideal*, meaning it has the tightest possible linear programming relaxation.

In particular, for each $\ell \in [\![|\mathbf{X}^k|]\!]$ and $s \in \mathcal{S}$, we apply the techniques of Balas to construct the strongest possible MILP formulation for the union of the two sets given as

$$\mathbf{P}_s^{1l} = \left\{ (\mathbf{t}_s, z_s, \boldsymbol{\tau}_s, \mathbf{y}_s) \in \mathbf{Q}_s(1) \;\middle|\; \begin{array}{l} \sum_{(i,j)\in\mathcal{A}} (c_{ij} + \beta\rho_{ij} + t_{sij}) x_{sij}^l \geq \\ \sum_{(i,j)\in\mathcal{A}} \Big[(c_{ij} + \beta\rho_{ij}) y_{sij}^{1l} + \tau_{sij}^{1l}\Big] + \epsilon_s \end{array} \right\},$$

$$\mathbf{P}_s^{2l} = \left\{ (\mathbf{t}_s, z_s, \boldsymbol{\tau}_s, \mathbf{y}_s) \in \mathbf{Q}_s(1) \;\middle|\; \sum_{(i,j)\in\mathcal{A}} n_s\Big(\rho_{ij} + \alpha(c_{ij} + t_{sij})\Big) x_{sij}^l - z_s \leq 0 \right\}$$

where

$$\mathbf{Q}_s(u) := \left\{ (\mathbf{t}_s, z_s, \boldsymbol{\tau}_s, \mathbf{y}_s) \left| \begin{array}{ll} \tau_{sij} \leq M^3_{sij} y_{sij} & \forall (i,j) \in \mathcal{A} \\ \tau_{sij} \leq t_{sij} & \forall (i,j) \in \mathcal{A} \\ \tau_{sij} \geq t_{sij} + M^3_{sij}(y_{sij} - u) & \forall (i,j) \in \mathcal{A} \\ \sum_{(i,j)\in\mathcal{A}} y_{sij} - \sum_{(j,i)\in\mathcal{A}} y_{sji} = b_{si} u & \forall i \in \mathcal{N} \\ \tau_{sij}, t_{sij} \geq 0 & \forall (i,j) \in \mathcal{A} \\ 0 \leq y_{sij} \leq u & \forall (i,j) \in \mathcal{A} \end{array} \right. \right\}. \tag{3.19}$$

Within the definition for $\mathbf{P}^{1l}_s$, $\mathbf{P}^{2l}_s$, and $\mathbf{Q}_s$, we take vectors defined as follows:

$$\begin{aligned} \mathbf{t}_s &= (t_{sij} : (i,j) \in \mathcal{A}) \\ \boldsymbol{\tau}_s &= (\tau_{sij} : (i,j) \in \mathcal{A}) \\ \mathbf{y}_s &= (y_{sij} : (i,j) \in \mathcal{A}). \end{aligned}$$

Observe that each constraint describing $\mathbf{Q}_s(u)$ is affine in u. Therfore, the following set is polyhedral.

$$\big\{(\mathbf{t}_s, z_s, \boldsymbol{\tau}_s, \mathbf{y}_s, u) | (\mathbf{t}_s, z_s, \boldsymbol{\tau}_s, \mathbf{y}_s) \in \mathbf{Q}_s(u)\big\}$$

Lemma 3.5. Select some $\ell \in [\![|\mathbf{X}^k|]\!]$ and $s \in \mathcal{S}$. The following is an ideal formulation for $\mathbf{P}^{1l}_s \cup \mathbf{P}^{2l}_s$:

$$\sum_{(i,j)\in\mathcal{A}} (c_{ij} u^{1l}_s + \beta \rho_{ij} u^{1l}_s + t^{1l}_{sij}) x^l_{sij} \geq \sum_{(i,j)\in\mathcal{A}} \Big[(c_{ij} + \beta\rho_{ij}) y^{1l}_{sij} + v^{1l}_{sij} \Big] + \epsilon_s u^{1l}_s \tag{3.20a}$$

$$\sum_{(i,j)\in\mathcal{A}} n_s \Big(\rho_{ij} u^{2l}_s + \alpha (c_{ij} u^{2l}_s + t^{2l}_{sij}) \Big) x^l_{sij} - z^{2l}_s \leq 0 \tag{3.20b}$$

$$(\mathbf{t}^{1l}_s, z^{1l}_s, \boldsymbol{\tau}^{1l}_s, \mathbf{y}^{1l}_s) \in \mathbf{Q}_s(u^{1l}_s) \tag{3.20c}$$

$$(\mathbf{t}_s^{2l}, z_s^{2l}, \boldsymbol{\tau}_s^{2l}, \mathbf{y}_s^{2l}) \in \mathbf{Q}_s(u_s^{1l}) \tag{3.20d}$$

$$(t_{sij}, z_s, \tau_{sij}, y_{sij}) = (t_{sij}^{1l}, z_s^{1l}, \tau_{sij}^{1l}, y_{sij}^{1l}) + (t_{sij}^{2l}, z_s^{2l}, \tau_{sij}^{2l}, y_{sij}^{2l}) \qquad \forall (i,j) \in \mathcal{A} \tag{3.20e}$$

$$u_s^{1l} + u_s^{2l} = 1 \tag{3.20f}$$

$$u_s^{1l}, u_s^{2l} \in \{0, 1\}. \tag{3.20g}$$

Applying the ideal MILP formulation (3.20) for every $l \in [\![|\mathbf{X}^k|]\!]$ and $s \in \mathcal{S}$, we can develop the following formulation for Master$_k$.

Proposition 3.6. The following is a valid MILP formulation for Master$_k$:

$$\text{minimize} \sum_{s \in \mathcal{S}} z_s \tag{3.21a}$$

subject to

$$\sum_{(i,j) \in \mathcal{A}} (c_{ij} u_s^{1l} + \beta \rho_{ij} u_s^{1l} + t_{sij}^{1l}) x_{sij}^l \geq \sum_{(i,j) \in \mathcal{A}} \Big[(c_{ij} + \beta \rho_{ij}) y_{sij}^{1l} + v_{sij}^{1l} \Big] + \epsilon_s u_s^{1l} \tag{3.21b}$$

$$\forall s \in \mathcal{S}, l \in [\![|\mathbf{X}^k|]\!]$$

$$\sum_{(i,j) \in \mathcal{A}} n_s \Big(\rho_{ij} u_s^{2l} + \alpha (c_{ij} u_s^{2l} + t_{sij}^{2l}) \Big) x_{sij}^l - z_s^{2l} \leq 0 \quad \forall s \in \mathcal{S}, l \in [\![|\mathbf{X}^k|]\!] \tag{3.21c}$$

$$(t_{sij}, z_s, \tau_{sij}, y_{sij}) = (t_{sij}^{1l}, z_s^{1l}, \tau_{sij}^{1l}, y_{sij}^{1l}) + (t_{sij}^{2l}, z_s^{2l}, \tau_{sij}^{2l}, y_{sij}^{2l}) \tag{3.21d}$$

$$\forall s \in \mathcal{S}, l \in [\![|\mathbf{X}^k|]\!], (i,j) \in \mathcal{A}$$

$$u_s^{1l} + u_s^{2l} = 1 \qquad \forall s \in \mathcal{S}, l \in [\![|\mathbf{X}^k|]\!] \tag{3.21e}$$

$$\tau_{sij}^{1l} \leq M_{sij}^3 y_{ij}^{1sl} \qquad \forall s \in \mathcal{S}, l \in [\![|\mathbf{X}^k|]\!], (i,j) \in \mathcal{A} \tag{3.21f}$$

$$\tau_{sij}^{2l} \leq M_{sij}^3 y_{ij}^{2sl} \qquad \forall s \in \mathcal{S}, l \in [\![|\mathbf{X}^k|]\!], (i,j) \in \mathcal{A} \tag{3.21g}$$

$$\tau_{sij}^{1l} \leq t_{sij}^{1l} \qquad \forall s \in \mathcal{S}, l \in [\![|\mathbf{X}^k|]\!], (i,j) \in \mathcal{A} \tag{3.21h}$$

$$\tau_{sij}^{2l} \leq t_{sij}^{2l} \qquad \forall s \in \mathcal{S}, l \in [\![|\mathbf{X}^k|]\!], (i,j) \in \mathcal{A} \tag{3.21i}$$

$$\tau_{sij}^{1l} \geq t_{sij}^{1l} + M_{sij}^3(y_{sij}^{1l} - u_s^{1l}) \qquad \forall s \in \mathcal{S}, l \in [\![|\mathbf{X}^k|]\!], (i,j) \in \mathcal{A} \tag{3.21j}$$

$$\tau_{sij}^{2l} \geq t_{sij}^{2l} + M_{sij}^3(y_{sij}^{2l} - u_s^{2l}) \qquad \forall s \in \mathcal{S}, l \in [\![|\mathbf{X}^k|]\!], (i,j) \in \mathcal{A} \tag{3.21k}$$

$$\sum_{(i,j)\in\mathcal{A}} y_{sij}^{1l} - \sum_{(j,i)\in\mathcal{A}} y_{sji}^{1l} = b_{si}u_s^{1l} \qquad \forall i \in \mathcal{N}, s \in \mathcal{S} \tag{3.21l}$$

$$\sum_{(i,j)\in\mathcal{A}} y_{sij}^{2l} - \sum_{(j,i)\in\mathcal{A}} y_{sji}^{2l} = b_{si}u_s^{2l} \qquad \forall i \in \mathcal{N}, s \in \mathcal{S} \tag{3.21m}$$

$$\tau_{sij}^{1l}, \tau_{sij}^{2l}, t_{sij}^{1l}, t_{sij}^{2l} \geq 0 \qquad \forall s \in \mathcal{S}, l \in [\![|\mathbf{X}^k|]\!], (i,j) \in \mathcal{A} \tag{3.21n}$$

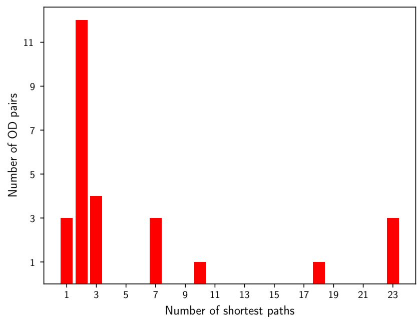
$$y_{sij}^{1l} \leq u^{1ls} \qquad \forall s \in \mathcal{S}, l \in [\![|\mathbf{X}^k|]\!], (i,j) \in \mathcal{A} \tag{3.21o}$$

$$y_{sij}^{2l} \leq u^{2ls} \qquad \forall s \in \mathcal{S}, l \in [\![|\mathbf{X}^k|]\!], (i,j) \in \mathcal{A} \tag{3.21p}$$

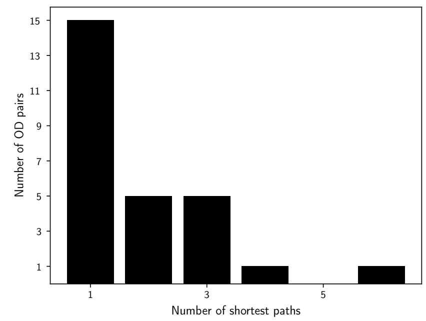
$$y_{sij}^{1l}, y_{sij}^{2l} \in \{0,1\} \qquad \forall s \in \mathcal{S}, l \in [\![|\mathbf{X}^k|]\!], (i,j) \in \mathcal{A} \tag{3.21q}$$

$$u_s^{1l}, u_s^{2l} \in \{0,1\} \qquad \forall s \in \mathcal{S}, l \in [\![|\mathbf{X}^k|]\!] \tag{3.21r}$$

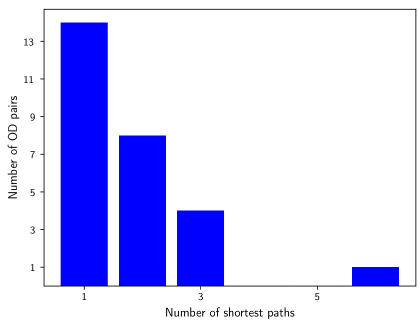
Note that the formulation above has more variables compare to the big-M formulation (3.16). However, constraints (3.21b) and (3.21c) have replaced big-M formulation constraints (3.16b) and (3.16c).



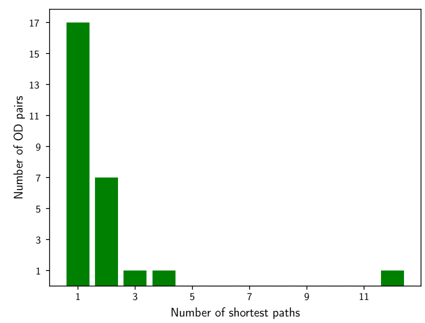
(a) $\alpha = 0, \beta = 0$



(b) $\alpha = 0, \beta = 1$



(c) $\alpha = 10^{-5}, \beta = 0$



(d) $\alpha = 10^{-5}, \beta = 1$

Figure 3.2 – Number of shortest paths for the Ravenna network under optimistic toll

3.6 Numerical Experiments

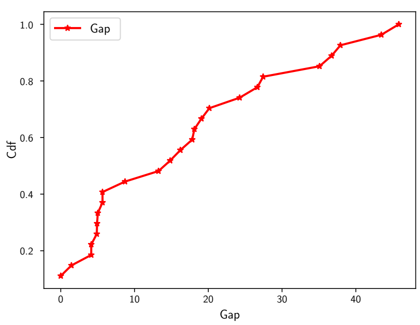
In this section, we present numerical experiments to test the validity of the proposed models and methods. We use the Julia Programming Language for modeling optimization problems, and we use Gurobi solver for solving MILPs.

3.6.1 Optimistic Approach Generates Multiple Lower-Level Optimal Solutions

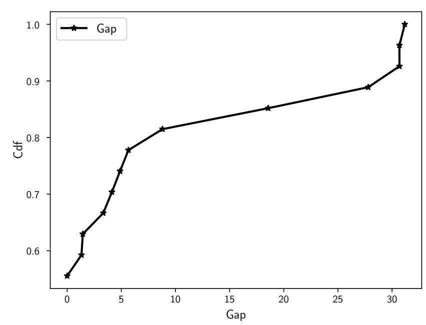
As described in Section 3.3, the lower level problem in HTP may have multiple solutions for the optimistic toll vector $\mathbf{t}^{\text{HTP}}$. To test the existence of multiple solutions for the lower level problem of HTP in relatively large networks, we consider the Ravenna network (Bonvicini and Spadoni, 2008). The Ravenna network has 111 nodes, 286 arcs, and 31 origin-destination pairs for four hazmat types: chlorine, LPG, gasoline, and methanol. Note that as the probability of the risk exposure is a very small number, to avoid the numerical issues rising from the existence of small numbers in our model, we normalize the probability vector by dividing its elements by its maximum value. We assume that the toll imposed on each arc is the same for each hazmat type. To find the solutions of the inner problem under $\mathbf{t}^{\text{HTP}}$, first we solve HTP, using the single level formulation (Marcotte et al., 2009). Next, we use Yen's algorithm for finding all the optimal solutions to the inner problem under optimistic toll vector $\mathbf{t}^{\text{HTP}}$. Figure 3.2a represents the histogram of the number of OD pairs based on the number of shortest paths for four different cases based on the values of α and β. In all cases, there are OD pairs with multiple shortest paths. As Figure 3.2a represents, the number of OD pairs with multiple shortest paths may be greater than the number of OD pairs with a single shortest path. Note that using a positive term for the risk in the lower level problem ($\beta > 0$) does not necessarily yield a single shortest path for each OD pair, as shown in Figure 3.2b and Figure 3.2d.

Next, we compare the upper-level objective function value for the best-case solution $\mathbf{x}^b \in \mathbb{X}(\mathbf{t}^{\text{HTP}})$ and the worst-case solution $\mathbf{x}^w \in \mathbb{X}(\mathbf{t}^{\text{HTP}})$ of the lower level problem under $\mathbf{t}^{\text{HTP}}$. Figure 3.3 represents the cumulative distribution of the gap between $f(\mathbf{t}^{\text{HTP}}, \mathbf{x}^b)$ and $f(\mathbf{t}^{\text{HTP}}, \mathbf{x}^w)$ for different values of α and β. As Figure 3.3 shows, the gap is greater than 25 % for some OD pairs for all cases. Moreover, more than 10% of the OD pairs have a gap greater than 15% for all cases. These results show the downside of the optimistic formulation HTP; the network users may not choose the best-case path, and as a result, they may face much higher risk than what is predicted by the optimistic approach.

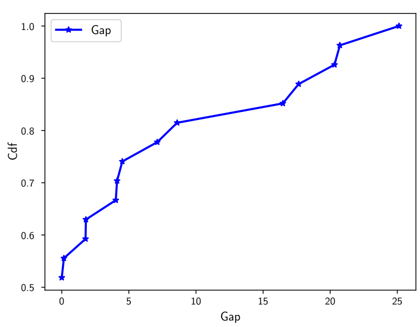
Next, to study the effect of toll policy on the network and the existence of multiple solutions in more detail, we consider the origin-destination pair $s = (109, 111)$ for the Ravenna network. Table 3.1 compares the toll free cost C_s, the total toll charged T_s, the total cost $C_s + T_s$, the upper-level objective function value f_s and the risk component of f_s denoted by R_s for the worst-case solution under $\mathbf{t}^{\text{HTP}}$ which is denoted by $\mathbf{x}_s^w \in \mathbb{X}_s(\mathbf{t}^{\text{HTP}})$, the best-case solution under $\mathbf{t}^{\text{HTP}}$ which is denoted by $\mathbf{x}_s^b \in \mathbb{X}_s(\mathbf{t}^{\text{HTP}})$ and the toll free solution $\mathbf{x}_s^0 \in \mathbb{X}_s(\mathbf{0})$



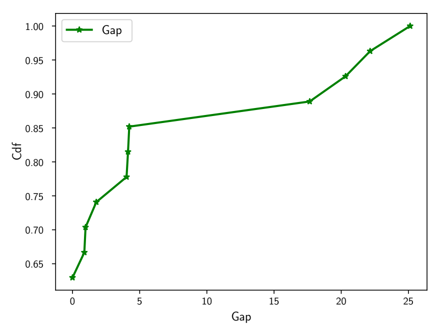
(a) $\alpha = 0, \beta = 0$



(b) $\alpha = 0, \beta = 1$



(c) $\alpha = 10^{-5}, \beta = 0$



(d) $\alpha = 10^{-5}, \beta = 1$

Figure 3.3 – CDF of the gap between $f(t^{\text{HTP}}, x^b)$ and $f(t^{\text{HTP}}, x^w)$ where $x^b, x^w \in \mathbb{X}(t^{\text{HTP}})$ for Ravenna network

Table 3.1 – The effect of toll policy for OD pair $s = (109, 111)$ for the Ravenna network ($\alpha = 10^{-5}, \beta = 0$)

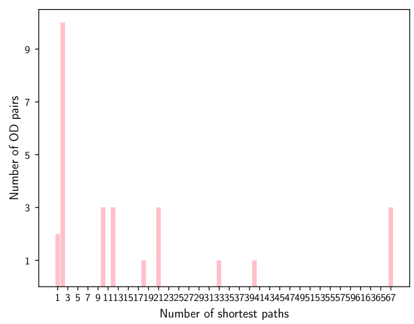
Solution	C_s	T_s	$C_s + T_s$	f_s	R_s
$\mathbf{x}_s^w$	29296.94	7598.78	36895.72	1423.73	1024.16
$\mathbf{x}_s^b$	36895.72	0	36895.72	1183.28	783.71
$\mathbf{x}_s^0$	25894.91	11000.81	36895.72	1392.17	992.59

when no road has been tolled. As presented in Table 3.1, $\mathbf{x}_s^0$ has the minimum toll free cost C_s. Under the optimistic toll policy, a relatively large toll T_s will be charged on path $\mathbf{x}_s^0$. As a result, other paths with lower risk become the shortest paths as well. Moreover, under the optimistic toll policy, there are multiple shortest paths (paths with the minimum total cost of $T_s + C_s$). As presented, the total cost $T_s + C_s$ is the same for all three paths, and all of them can be chosen by network users in the lower level problem. However, the upper-level objective function f_s is different, and it is possible that the network users choose $\mathbf{x}_s^w(\mathbf{t}^{\text{HTP}})$, which has the highest risk.

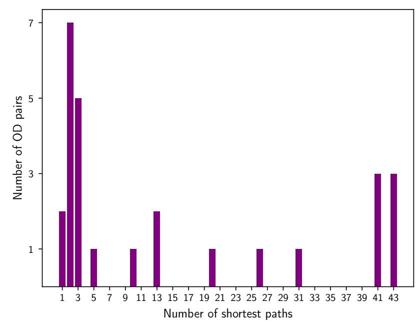
Next, we consider the optimistic approach in the context of boundedly rational behavior where users do not necessarily choose the shortest path.

3.6.2 Optimistic Approach Generates Multiple Lower-Level Satisficing Solutions

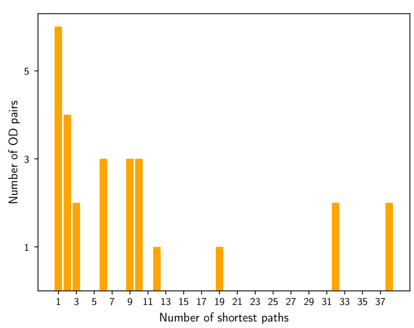
We also examine the set of satisficing solutions. As the set of shortest paths is a subset of ϵ-approximate paths, the set of solutions to the inner problem is larger, and the gap between the objective function value for the worst-case and the best-case solution is greater under satisficing network users. Figure 3.4 represents the histogram of the number of OD pairs based on the number of satisficing paths for different values of α and β when the the indifference band is set to 10 % of the length of the shortest path when no toll has been charged ($\epsilon_s = 0.1 \min_{\mathbf{y'} \in \mathbf{X}} h_s(\mathbf{0}, \mathbf{y'})$). As Figure 3.4 shows, for the Ravenna network, there may be many ϵ-approximate paths under optimistic toll vector $\mathbf{t}^{\text{HTP}}$. For instance, as Figure 3.4a represents, the number of shortest paths for an OD pair can be as high as 67. Similar to the case of shortest paths, using a positive term for the risk ($\beta > 0$) in lower level problem



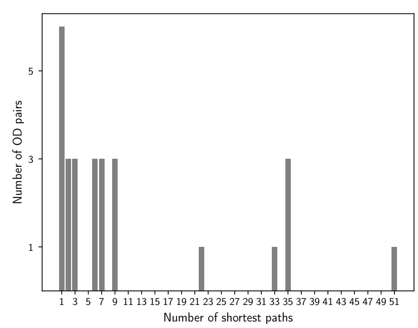
(a) $\alpha = 0, \beta = 0$



(b) $\alpha = 0, \beta = 1$

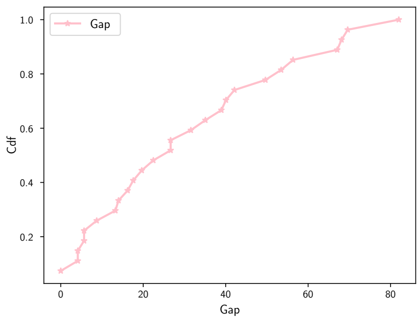


(c) $\alpha = 10^{-5}, \beta = 0$

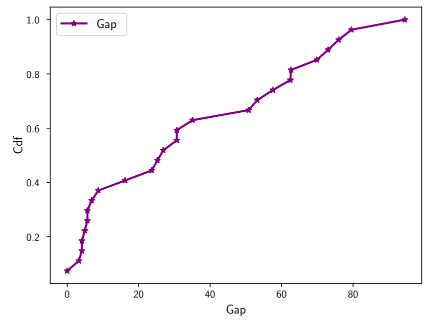


(d) $\alpha = 10^{-5}, \beta = 1$

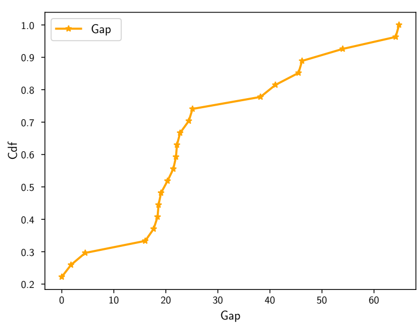
Figure 3.4 – Number of ϵ-approximate paths for the Ravenna network



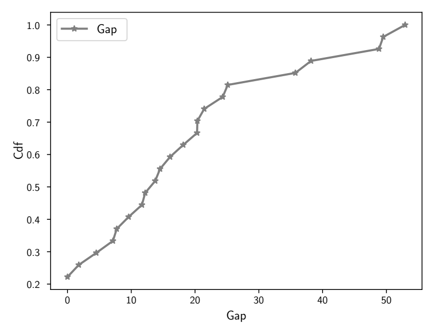
(a) $\alpha = 0, \beta = 0$



(b) $\alpha = 0, \beta = 1$



(c) $\alpha = 10^{-5}, \beta = 0$



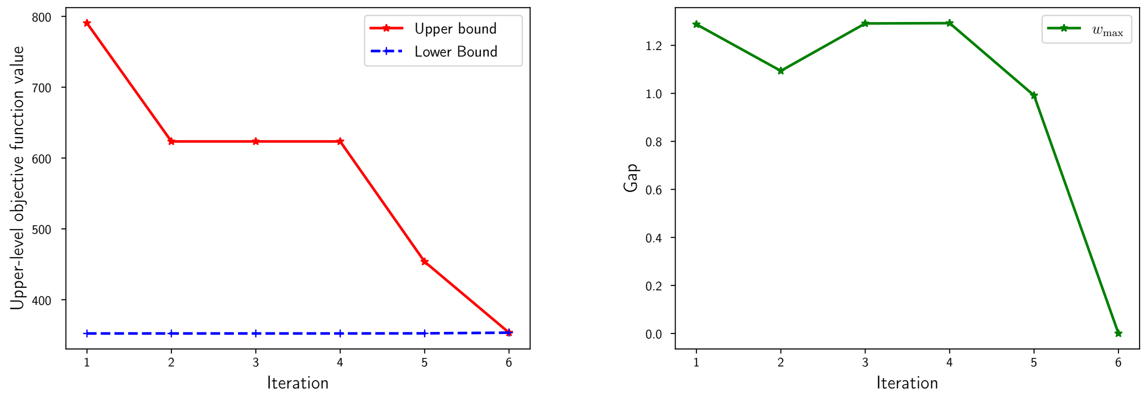
(d) $\alpha = 10^{-5}, \beta = 1$

Figure 3.5 – CDF of the gap between $f(t^{\text{HTP}}, x^b)$ and $f(t^{\text{HTP}}, x^w)$ where $x^b, x^w \in \mathbb{X}^\epsilon(t^{\text{HTP}})$

does not necessarily yield to single boundedly rational path for each OD pair as Figure 3.4b and Figure 3.4d represent. Figure 3.5 represents the cumulative distribution of the gap between $f(\mathbf{t}^{\text{HTP}}, \mathbf{x}^b)$ and $f(\mathbf{t}^{\text{HTP}}, \mathbf{x}^w)$ for different values of α and β under $\mathbf{t}^{\text{HTP}}$. Note that here $\mathbf{x}^b, \mathbf{x}^w \in \mathbb{X}^\epsilon(\mathbf{t}^{\text{HTP}})$. As Figure 3.5 shows, the gap is greater when network users are satisficing and applying optimistic approach may result in a much higher risk from what is predicted by optimistic approach.

3.6.3 Cutting Plane Algorithm for Solving Approximate Pessimistic Problem

In order to implement the cutting plane algorithm in Section 3.5, we add the valid inequality described in Section 3.5.2 to Master$_k$ and use the results in Section 3.6 for bounding big-M constants.



(a) Lower bound and upper bound

(b) $w_{\max}$

Figure 3.6 – Cutting plane algorithm for solving approximate pessimistic problem for the Nine-node network

We implement the exact cutting plane algorithm in Section 3.5.1 for the Nine-node network (Hearn and Ramana, 1998). The Nine-node network includes 9 nodes and 18 links. We consider 16 origin-destination pairs. We consider the free flow time of each link as the cost associated with that link. For experimental purposes, we assume that n_s is a discrete random number between 0 and 100, and ρ_{ij}^s is a random number between 0 and 1. We set the value of α and β to 10^{-2}. We also set the value of $\epsilon_s = 0.1 \min_{\mathbf{x}\in\mathbf{X}} h_s(\mathbf{0}, \mathbf{x})$ for each shipment $s \in \mathcal{S}$.

Figure 3.6a represents the lower bound and the upper bound for $f(\cdot)$ in each iteration of the cutting plane algorithm. Figure 3.6b denote $\max_{s\in\mathcal{S}} w_s$ in sub-problem Sub_k. As figures 3.6a and 3.6b present, the algorithm terminates after 6 iterations. In the last iteration, the lower bound and the upper bound have the same value and $w_s \leq 0$ for all $s \in \mathcal{S}$.

As mentioned earlier, for solving Master_k, we can either use big-M formulation or Balas formulation. Although theoretically interesting, since Balas formulation has many variables and constraints, it may not be a suitable approach for solving Master_k for large networks. Table 3.2 compares Balas formulation and big-M formulation for the Nine-node network. We report the number of variables and constraints for Master_k in the last round of the cutting plane algorithm, the total number of nodes explored for solving Master_k and the

Table 3.2 – Comparing Balas and Big-M formulation for Master_k in the final iteration

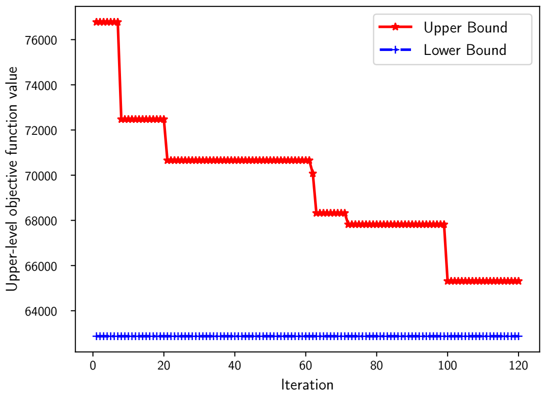
	# of constraints	# of variables	Nodes explored	Time(s)
Balas	21216	11266	13875	74.28
Big-M	1216	706	5567	2.25

total solution time of Master_k for all the iterations of the algorithm. As it can be seen, big-M formulation has fewer variables and constraints and is much faster compare to the Balas formulation.

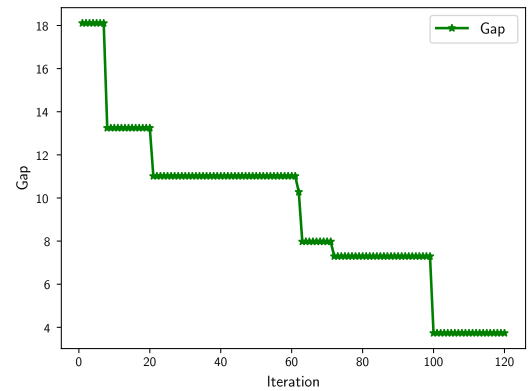
Next, we test the performance of the cutting plane algorithm for the Ravenna network. We set the value of $\alpha = 10^{-5}$ and $\beta = 0$. We also set the value of $\epsilon_s = 0.1 \min_{\mathbf{x}} h_s(\mathbf{0}, \mathbf{x})$ for each shipment s. Note that we solve the master problem to 1 percent optimality gap in each iteration. We set the lower bound in iteration k as the highest lower bound obtained by solving the master problem until iteration k. Similarly, the upper bound is the lowest upper bound obtained until iteration k. To speed up the algorithm, we add multiple paths to the master problem in each iteration of the algorithm. Figure 3.7a shows the upper bound and the lower bound for PHTP_ϵ optimal objective value obtained by the cutting plane heuristic. The upper bound is non-increasing, and the lower bound is non-decreasing. As Figure 3.7 represents, while the algorithm starts with a gap of 18 percent, it could solve the problem with an optimality gap of 3.7 percent in 48 hours.

Note that the value of M^3_{sij} is usually a very large number. This may result in charging a high toll on some arcs, which may not be acceptable by network users. Additionally, setting a large number to M^3_{ij} results in numerical issues and makes the problem difficult to solve too. Considering this, we may consider a maximum toll amount that can be charged on a link by considering the following constraint:

$$t_{ij} \leq t^{\max}_{ij} \quad \forall (i,j) \in \mathcal{A} \tag{3.22}$$



(a) Lower bound and upper bound



(b) Optimality gap

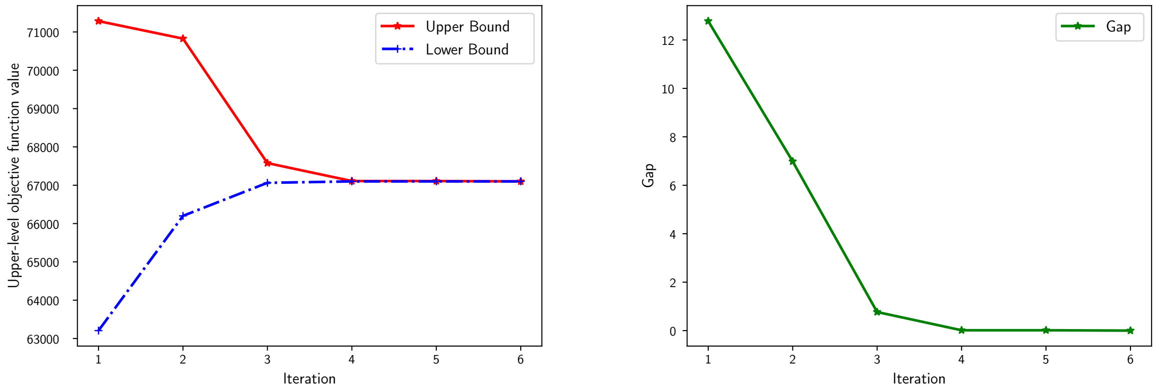
Figure 3.7 – Cutting plane algorithm for solving approximate pessimistic problem for the Ravenna network

Setting a bound on the maximum toll on each link helps us bound M_s^{1l}, M_s^{2l} and M_{sij}^3 in Master$_k$ easily. Figure 3.8a represents the upper bound and the lower bound for the cutting plane algorithm when the toll in each arc is bounded ($\mathbf{t}^{\max} = \mathbf{500}$). Similar to the unbounded toll case, the upper bound is non-increasing, and the lower bound is non-decreasing. As Figure 3.8b represents, we could solve the problem to optimality after 6 iterations.

Table 3.3 compares the PHTP$_\epsilon$ objective value when the toll on each link is upper bounded by $\mathbf{t}^{\max}$ which is denoted by $z^*_{\mathbf{t}\leq\mathbf{t}^{\max}}$ and when it is not bounded which is denoted by z^*. For the bounded case we consider solution time limit of 2 hours. As it can be seen when we increase $\mathbf{t}^{\max}$ the gap between $z^*_{\mathbf{t}\leq\mathbf{t}^{\max}}$ and z^* becomes smaller. In particular, for $\mathbf{t}^{\max} = \mathbf{1000}$ and $\mathbf{t}^{\max} = \mathbf{1500}$, bounding $\mathbf{t}$ by $\mathbf{t}^{\max}$ results in a better solution. The reason is that when we bound $\mathbf{t}$ by $\mathbf{t}^{\max}$, the feasible region of PHTP$_\epsilon$ becomes much tighter, and PHTP$_\epsilon$ becomes much easier to solve and we obtain a high-quality solution.

Table 3.3 – Comparing z^* and $z^*_{t\leq t^{\max}}$

$\mathbf{t}^{\max}$	$z^*_{\mathbf{t}\leq\mathbf{t}^{\max}}$	z^*	Gap (%)
500	67099.75	65310.98	2.66
1000	65063.24	65310.98	-0.35
1500	64930.22	65310.98	-0.58

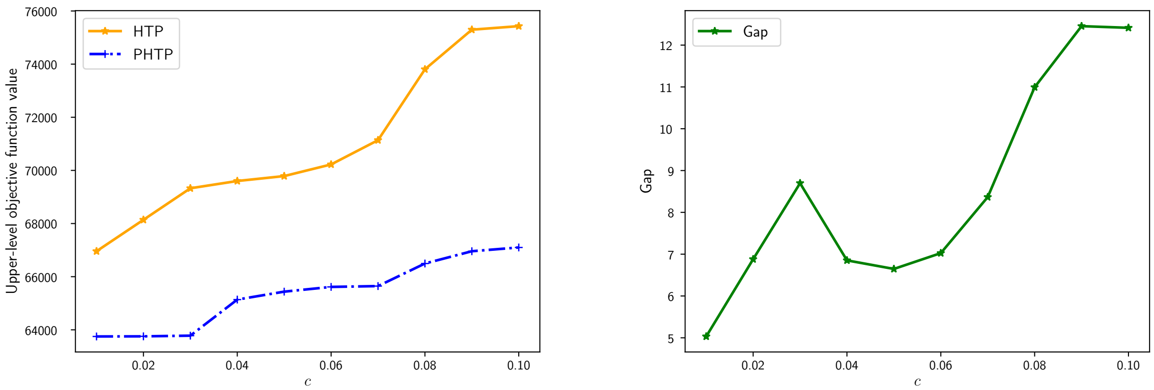


(a) Lower bound and upper bound (b) Optimality gap

Figure 3.8 – Cutting plane algorithm for solving PHTP_ϵ for the Ravenna network ($t^{\max} = 500$)

3.6.4 Comparing Optimistic and Pessimistic Approaches

In this section, we compare the optimistic and pessimistic policies. First, we solve the optimistic formulation HTP using single level formulation proposed by Marcotte et al. (2009). We compute the worst-case solution under optimistic toll vector $\mathbf{t}^{\text{HTP}}$ for different values of ϵ_s. Next, we solve PHTP_ϵ for different values of ϵ_s and compares it to the worst-case solution under $\mathbf{t}^{\text{HTP}}$. Figure 3.9a represents the worst-case upper level objective function under optimistic and pessimistic toll vectors. We denote this values by $f(\mathbf{x}^w, \mathbf{t}^{\text{HTP}})$ and $f(\mathbf{x}^w, \mathbf{t}^{\text{PHTP}_\epsilon})$. We represent $f(\mathbf{x}^w, \mathbf{t}^{\text{HTP}})$ and $f(\mathbf{x}^w, \mathbf{t}^{\text{PHTP}_\epsilon})$ for different values of constant c where $\epsilon_s = c \min_{\mathbf{x}} h_s(\mathbf{0}, \mathbf{x})$. We set $t_{ij}^{\max} = 500$ for all $(i,j) \in \mathcal{A}$. Note that $f(\mathbf{x}^w, \mathbf{t}^{\text{PHTP}_\epsilon})$ is always smaller than $f(\mathbf{x}^w, \mathbf{t}^{\text{HTP}})$. Even for small value of $c = 0.01$, there is 5 percent gap between $f(\mathbf{x}^w, \mathbf{t}^{\text{HTP}})$ and $f(\mathbf{x}^w, \mathbf{t}^{\text{PHTP}_\epsilon})$. Note that in general, as e increases, the indifference band becomes larger and the difference between PHTP_ϵ and the worst-case solution under $\mathbf{t}^{\text{HTP}}$ becomes larger.



(a) Upper level objective function value

(b) Gap

Figure 3.9 – Comparing the worst-case solution under optimistic toll and the pessimistic toll

Chapter 4: On the Price of Satisficing in Network Equilibria

The copyright permissions for reuse previously published material in this chapter can be found in Appendix A.

4.1 Introduction

Instead of assuming a perfectly rational person with a clear system of preferences and perfect knowledge of the surrounding decision-making environment, we can consider *boundedly* rational persons with (1) an ambiguous system of preferences and (2) lack of complete information, following Simon (1955). When decision makers are indifferent among alternatives within a certain threshold, they are called *satisficing* decision makers, opposed to *optimizing* decision makers. The notion of satisficing was first introduced by Simon (1955, 1956). Satisficing decision makers choose any alternative whose utility level is above a threshold, called an *aspiration level*, even when the alternative is not optimal. The satisficing behavior is related to the first source of boundedness—an ambiguous system of preferences.

In transportation research, modeling drivers' route choice is an important task. While the travel-time minimization has been traditionally used as a basis for such modeling, sub-optimal route-choice behavior has gained attention. Since Mahmassani and Chang (1987), bounded rationality has gained attention in the transportation research literature (Szeto and Lo, 2006; Wu et al., 2013; Han et al., 2015; Szeto and Lo, 2006; Ge and Zhou, 2012; Di et al., 2014; Guo, 2013; Lou et al., 2010). Empirical evidence supports bounded rationality of drivers (Nakayama et al., 2001; Zhu and Levinson, 2010). The notion of bounded rationality has also been considered in the evaluation of value of times in connection to route-choice modeling (Xu et al., 2017), and in the model of behavior adjustment process (Ye and Yang, 2017). We refer readers to a review of Di and Liu (2016). In the non-transportation literature, the notion of bounded rationality and satisficing has also received much attention (Charnes and Cooper, 1963; Lam et al., 2013; Jaillet et al., 2016; Chen et al., 1997; Brown and Sim, 2009).

While the above-mentioned transportation research literature considers boundedly rational drivers, their discussion is limited to satisficing drivers without considering the second source of boundedness: lack of complete information on the decision environment. Sun et al. (2018) connect the first and the second sources of boundedness by considering both satisficing behavior and incomplete information, in the context of shortest-path finding in *congestion-free* networks. Sun et al. (2018) study the second source by considering errors in drivers' perception of arc travel time, and conclude that their perception-error model can generally capture both sources of boundedness in rationality in a single unified modeling framework.

In the literature, the traditional network user equilibrium, Wardrop equilibrium in particular, is called the perfectly rational user equilibrium (PRUE), while a traffic pattern equilibrated among satisficing drivers is called a boundedly rational user equilibrium (BRUE). In this study, we will use a new term *satisficing user equilibrium* (SatUE) instead of BRUE to emphasize that it only considers the first source of boundedness without considering drivers' incomplete information on the decision environment. We believe that the term 'BRUE' should be used to describe a broader and more general class of models, including SatUE.

Note that SatUE differs from the stochastic user equilibrium (SUE) (Sheffi, 1985) in two important aspects. First, drivers are assumed to be optimizing decision makers in SUE, while they are satisficing in SatUE. Second, with appropriate probability distributions assumed in the random utility model in SUE, each path possesses a probability of being chosen; hence we can compute the expected traffic flow rate in each path. In SatUE, however, each satisficing path is acceptable to drivers, but it may or may not be chosen by drivers and we do not know its probability of being chosen. See further discussion in Di and Liu (2016).

The main contribution of this study is the quantification of how bad the total travel time in SatUE can be. In a SatUE traffic pattern, the total travel time can be either greater than or less than that of PRUE. We define the *price of satisficing* (PoSat) as the ratio between the worst-case total travel time of SatUE and the total travel time of PRUE. This research quantifies PoSat analytically and compares with numerical bounds.

The analytical quantification of PoSat is related to the price of anarchy (PoA) (Koutsoupias and Papadimitriou, 1999; Roughgarden and Tardos, 2002) that compares the performances of the system optimal solutions and the PRUE solutions. Using a similar idea, we can also compare the performance of the perfectly rational user equilibrium traffic patterns and satisficing user equilibrium traffic patterns. While PoA quantifies how much system-wide performance we can lose by competing, PoSat quantifies how much we can lose by satisficing. Roughgarden and Tardos (2002) define and study the PoA of *approximate Nash equilibria*, which are essentially SatUE patterns. We develop our bounds for PoSat based on the bounds for PoA of approximate Nash equilibria (Christodoulou et al., 2011) and the ideas from the sensitivity analysis of traffic equilibria (Dafermos and Nagurney, 1984). Note that Perakis (2007) studies the PoA of the *exact* Nash equilibria with general nonlinear, asymmetric cost functions.

The notion of PoSat is also related to the *price of risk aversion* (Nikolova and Stier-Moses, 2015) and the *deviation ratio* (Kleer and Schäfer, 2016). When network users are risk-averse decision makers, the price of risk aversion compares the performances of the resulting equilibrium among risk-averse users and the (risk-neutral) PRUE. When network users' cost functions are deviated from the true cost functions for some reasons, the deviation ratio compares the performances of the resulting equilibrium and the PRUE. Kleer and Schäfer (2016) show that the price of risk aversion is a special case of the deviation ratio. In both research articles, however, only cases with a common single origin node are considered. In this study, we consider general cases with multiple origin nodes and multiple destination nodes, with asymmetric travel time functions.

This chapter is organized as follows. In Section 4.2, we introduce the notation and define various concepts including user equilibrium, system optimum, satisficing behavior, price of anarchy, and price of satisficing. In Section 4.3, we define the user equilibrium with perception errors and make connections with satisficing user equilibrium. Our main result is introduced in Section 4.4, where we derive the analytical worst-case bound on the price of satisficing. In Section 4.5, we compare the analytical bound with numerical bounds.

4.2 Notation and Definitions

Since we will use path-based and arc-based flow variables and their corresponding functions and sets interchangeably, we need clear definitions of variables, sets, and functions. We use boldfaced lower-case letters for vector quantities as in $\mathbf{v}$ and normal lower-case letters for their components as in v_a; similarly, vector-valued functions like $\mathbf{t}(\cdot)$ and their components like $t_a(\cdot)$. We use boldfaced upper-case letters for the set that they belong to, as in $\mathbf{v} \in \mathbf{V}$. We use calligraphic capital letters for sets of indices as in $\mathcal{N}$. The only exception is that $\mathbf{Q}$ (a bold-face capital letter instead of lower case) represents a vector of Q_w, the demand for OD pair w. The lower-case version q_i^w is instead the net amount of flow associated with OD pair w that enters or leaves node i in the next subsection

4.2.1 Traffic Flow Variables and Feasible Sets

We consider a network with a set of origin and destination $\mathcal{W}$ that is represented by directed graph $G(\mathcal{N}, \mathcal{A})$, where $\mathcal{N}$ is the set of nodes, and $\mathcal{A}$ is the set of arcs. For each origin destination pair $w \in \mathcal{W}$, the demand is Q_w and the set of available paths is $\mathcal{P}_w$. The set of all available paths in the whole network is defined as $\mathcal{P} = \cup_{w \in \mathcal{W}} \mathcal{P}_w$.

We also define the set of path flow variables $\mathbf{f}$ as

$$\mathbf{F} = \left\{ \mathbf{f} : \sum_{p \in \mathcal{P}_w} f_p = Q_w \quad \forall w \in \mathcal{W}, \qquad f_p \geq 0 \quad \forall p \in \mathcal{P} \right\}$$

and the corresponding set of arc flow variables $\mathbf{v}$ is defined as

$$\mathbf{V} = \left\{ \mathbf{v} : v_a = \sum_{p \in \mathcal{P}} \delta_a^p f_p \quad \forall a \in \mathcal{A}, \qquad \mathbf{f} \in \mathbf{F} \right\}$$

where $\delta_a^p = 1$ if path p contains arc a and $\delta_a^p = 0$ otherwise. Let $\mathcal{A}_i^+$ and $\mathcal{A}_i^-$ be the set of arcs whose tail node and head node are i, respectively. When we need to preserve OD information in arc flow variables, we use $\mathbf{x}$ as follows:

$$\begin{aligned} \mathbf{X} &= \left\{ \mathbf{x} : x_a^w = \sum_{p \in \mathcal{P}_w} \delta_a^p f_p \quad \forall a \in \mathcal{A}, w \in \mathcal{W} \qquad \mathbf{f} \in \mathbf{F} \right\} \\ &= \left\{ \mathbf{x} : \sum_{a \in \mathcal{A}_i^+} x_a^w - \sum_{a \in \mathcal{A}_i^-} x_a^w = q_i^w \quad \forall w \in \mathcal{W}, i \in \mathcal{N} \right\} \end{aligned}$$

where $q_i^w = -Q_w$ if $i = o(w)$, $q_i^w = Q_w$ if $i = d(w)$, and $q_i^w = 0$ otherwise.

We have $v_a = \sum_{p \in \mathcal{P}} \delta_a^p f_p$, $x_a^w = \sum_{p \in \mathcal{P}_w} \delta_a^p f_p$, and $v_a = \sum_{w \in \mathcal{W}} x_a^w$. Therefore, the transformations from $\mathbf{f}$ to $\mathbf{v}$, from $\mathbf{f}$ to $\mathbf{x}$, and from $\mathbf{x}$ to $\mathbf{v}$ are unique, which are denoted by $\mathbf{f} \mapsto \mathbf{v}$, $\mathbf{f} \mapsto \mathbf{x}$, and $\mathbf{x} \mapsto \mathbf{v}$, respectively. The inverse transformations are, however, not unique. In the rest of this chapter, to emphasize the non-uniqueness of the transformation and refer to *any* result of such transformation, we use $\overset{\text{any}}{\longmapsto}$; for example, with $\mathbf{v} \overset{\text{any}}{\longmapsto} \mathbf{f}$, we consider any $\mathbf{f}$ such that $v_a = \sum_{p \in \mathcal{P}} \delta_a^p f_p$.

We will use $\mathbf{v}$, $\mathbf{f}$, and $\mathbf{x}$ interchangeably to describe the same traffic pattern. In particular, we define

- $\mathbf{f}^*$, $\mathbf{v}^*$, $\mathbf{x}^*$: system optimal flow vectors (Section 4.2.2)
- $\mathbf{f}^0$, $\mathbf{v}^0$, $\mathbf{x}^0$: perfectly rational user equilibrium flow vectors (Section 4.2.3)
- $\mathbf{f}^\kappa$, $\mathbf{v}^\kappa$, $\mathbf{x}^\kappa$: (multiplicative) satisficing user equilibrium flow vectors with a multiplicative factor (to be defined subsequently) κ (Section 4.2.4)

Note that when $\kappa = 0$, we have $\mathbf{f}^\kappa = \mathbf{f}^0$.

4.2.2 Travel Time Functions and System Optimum

We denote arc travel function with arc traffic volume $\mathbf{v}$ by $t_a(\mathbf{v})$ for each arc $a \in \mathcal{A}$. We consider a performance function for each arc a as

$$z_a(\mathbf{v}) = t_a(\mathbf{v})v_a.$$

We denote the travel time function along path p with flow $\mathbf{f}$ by $c_p(\mathbf{f})$. When written as functions of $\mathbf{x}$, the arc travel time is denoted as $\tau_a(\mathbf{x})$ or $\tau_a^w(\mathbf{x})$, where the latter is used to emphasize the focus on OD pair w. Of course, $\tau_a^w(\mathbf{x}) = \tau_a(\mathbf{x}) = t_a(\mathbf{v})$, where $\mathbf{v} = \sum_w x_a^w$. The performance function for path $p \in \mathcal{P}$ is as follows:

$$z_p(\mathbf{f}) = c_p(\mathbf{f})f_p.$$

The following shows the relationship between path and arc travel times.

$$c_p(\mathbf{f}) = \sum_{a \in \mathcal{A}} \delta_a^p t_a(\mathbf{v}).$$

We define the arc-based total system performance function $Z(\mathbf{v})$ and path-based total system performance function $C(\mathbf{f})$ interchangeably as follows:

$$\begin{aligned} Z(\mathbf{v}) &\equiv \sum_{a \in \mathcal{A}} z_a(\mathbf{v}) = \sum_{a \in \mathcal{A}} t_a(\mathbf{v})v_a \\ &= \sum_{p \in \mathcal{P}} z_p(\mathbf{f}) = \sum_{p \in \mathcal{P}} c_p(\mathbf{f})f_p = \sum_{w \in \mathcal{W}} \sum_{p \in \mathcal{P}_w} c_p(\mathbf{f})f_p \equiv C(\mathbf{f}), \end{aligned}$$

which is also called the total travel time. A flow pattern that minimizes $Z(\cdot)$ or $C(\cdot)$ is called a *system optimal* flow pattern.

The vector-valued function $\mathbf{t}(\cdot)$ is called *monotone* in $\mathbf{V}$ if

$$[\mathbf{t}(\mathbf{v}^1) - \mathbf{t}(\mathbf{v}^2)]^\top (\mathbf{v}^1 - \mathbf{v}^2) \geq 0 \tag{4.1}$$

for all $\mathbf{v}^1, \mathbf{v}^2 \in \mathbf{V}$. If (4.1) holds as a strict inequality for all $\mathbf{v}^1 \neq \mathbf{v}^2$, it is said *strictly monotone*. The function $\mathbf{t}(\cdot)$ is called *strongly monotone* in $\mathbf{V}$ with modulus $\alpha > 0$ if

$$[\mathbf{t}(\mathbf{v}^1) - \mathbf{t}(\mathbf{v}^2)]^\top (\mathbf{v}^1 - \mathbf{v}^2) \geq \alpha \left\|\mathbf{v}^1 - \mathbf{v}^2\right\|_{\mathbf{V}}^2 \tag{4.2}$$

for all $\mathbf{v}^1, \mathbf{v}^2 \in \mathbf{V}$, where $\|\cdot\|_{\mathbf{V}}$ is the l^2-norm in $\mathbf{V}$. The monotonicity of path-based travel time function $c_p(\cdot)$ or its vector form $\mathbf{c}(\cdot)$ can be similarly defined. The path-based function $c_p(\cdot)$, however, is not strongly monotone in general (e.g., see Example 3 in de Palma and Nesterov, 1998).

4.2.3 Perfectly Rational User Equilibrium

When network users are perfectly rational—they seek the shortest path—we attain the perfectly rational user equilibrium (PRUE) defined as follows:

Definition 1 (Perfectly Rational User Equilibrium)**.** A traffic pattern $\mathbf{f}^0$ is called a *perfectly rational user equilibrium* (PRUE), if

$$\text{(PRUE)} \qquad f_p^0 > 0 \implies c_p(\mathbf{f}^0) = \min_{p' \in \mathcal{P}_w} c_{p'}(\mathbf{f}^0) \tag{4.3}$$

for all $p \in \mathcal{P}_w$ and $w \in \mathcal{W}$.

Using the arc travel function, the above condition can be restated as follows

$$f_p^0 > 0 \implies \sum_{a \in \mathcal{A}} \delta_a^p t_a(\mathbf{v}^0) = \min_{p' \in \mathcal{P}_w} \sum_{a \in \mathcal{A}} \delta_a^{p'} t_a(\mathbf{v}^0) \tag{4.4}$$

for all $p \in \mathcal{P}_w$ and $w \in \mathcal{W}$.

It is well known that a solution to the following variational inequality problem is a user equilibrium traffic flow (Smith, 1979; Dafermos, 1980):

$$\text{to find } \overline{\mathbf{f}} \in \mathbf{F} : \sum_{p \in \mathcal{P}} c_p(\overline{\mathbf{f}})(f_p - \overline{f}_p) \geq 0 \qquad \forall \mathbf{f} \in \mathbf{F}, \tag{4.5}$$

which can be equivalently rewritten as:

$$\text{to find } \overline{\mathbf{v}} \in \mathbf{V} : \sum_{a \in \mathcal{A}} t_a(\overline{\mathbf{v}})(v_a - \overline{v}_a) \geq 0 \qquad \forall \mathbf{v} \in \mathbf{V}, \tag{4.6}$$

or

$$\text{to find } \overline{\mathbf{x}} \in \mathbf{X} : \sum_{a \in \mathcal{A}} \sum_{w \in \mathcal{W}} \tau_a(\overline{\mathbf{x}})(x_a^w - \overline{x}_a^w) \geq 0 \qquad \forall \mathbf{x} \in \mathbf{X} \tag{4.7}$$

where $\tau_a^w(\mathbf{x}) = \tau_a(\mathbf{x}) = t_a(\mathbf{v})$.

With strictly monotone functions $t_a(\cdot)$, the solution $\overline{\mathbf{v}}$ to (4.6) is unique. While the transformations $\overline{\mathbf{v}} \overset{\text{any}}{\longmapsto} \overline{\mathbf{f}}$ and $\overline{\mathbf{v}} \overset{\text{any}}{\longmapsto} \overline{\mathbf{x}}$ are not unique, any such $\overline{\mathbf{f}}$ and $\overline{\mathbf{x}}$ are solutions to (4.5) and (4.7), respectively; therefore, solutions to (4.5) and (4.7) are not unique in general.

When the travel time on arc a is a function of only v_a, i.e. $t_a = t_a(v_a)$, then it is called *separable*. With separable arc travel time functions, the variational inequality problem (4.6) admits an equivalent convex optimization problem as formulated by (Beckmann et al., 1956). In general, if the Jacobian matrix of the arc travel time function vector $\mathbf{t}(\mathbf{v})$ is symmetric, that is,

$$\frac{\partial t_a(\mathbf{v})}{\partial v_e} = \frac{\partial t_e(\mathbf{v})}{\partial v_a} \qquad \forall a, e \in \mathcal{A},$$

for all $\mathbf{v} \in \mathbf{V}$, the variational inequality problem (4.6) can be reformulated as an equivalent Beckmann-type convex optimization problem (Patriksson, 2015; Friesz and Bernstein, 2016). When the Jacobian is asymmetric, no Beckmann-type convex optimization problem equivalent to (4.6) exists in general. In this case, the arc travel time functions is characterized as asymmetric and obtaining a PRUE flow requires solving a variational inequality problem.

4.2.4 Satisficing User Equilibrium

We introduce definitions of satisficing behavior and corresponding user equilibrium traffic patterns. In transportation research literature, boundedly rational user equilibrium (BRUE) is often defined with an additive term (see e.g., Lou et al., 2010; Di et al., 2013; Han et al., 2015). Herein, we refer to BRUE in the literature as an 'additive satisficing user equilibrium' to (i) highlight its additive feature and (ii) limit user behavior to just satisficing. Bounded rationality includes behaviors other than satisficing as well.

Definition 2 (Additive Satisficing)**.** A traffic pattern $\mathbf{f}$ is called an *additive satisficing user equilibrium (ASatUE)* with an additive factor E , if

$$\text{(ASatUE)} \qquad f_p > 0 \implies c_p(\mathbf{f}) \leq \min_{p' \in \mathcal{P}_w} c_{p'}(\mathbf{f}) + E \tag{4.8}$$

for all $p \in \mathcal{P}_w$ and $w \in \mathcal{W}$, where E is a positive constant.

We can also derive a similar definition using a multiplicative term. While the additive form in Definition 2 is popularly used in the transportation research literature, the multiplicative form in Definition 3 enables us to consider the satisficing level in disaggregate arc levels as we will observe in this research. Multiplicative satisficing user equilibrium is also called approximate Nash equilibrium in the price of anarchy literature (Christodoulou et al., 2011).

Definition 3 (Multiplicative Satisficing)**.** A traffic pattern $\mathbf{f}^\kappa$ is called a *multiplicative satisficing user equilibrium* with a multiplicative factor κ, or κ-MSatUE, if

$$\text{(MSatUE)} \qquad f_p^\kappa > 0 \implies c_p(\mathbf{f}^\kappa) \leq (1+\kappa) \min_{p' \in \mathcal{P}_w} c_{p'}(\mathbf{f}^\kappa) \tag{4.9}$$

for all $p \in \mathcal{P}_w$ and $w \in \mathcal{W}$, where $\kappa \geq 0$ is a constant.

Note that the additive (E) and multiplicative (κ) factor in (4.8) and (4.9), respectively, may be defined for each OD pair w. For example, E_w and κ_w can replace E and κ in (4.8) and (4.9), respectively, to allow for non-homogeneous satisficing thresholds. In such cases, however, we assume that travelers for the same OD pair are homogeneous with the same threshold E_w or κ_w. In this research, to describe the satisficing behavior, we focus only on MSatUE. Moreover, for simplicity, we use a single value of κ for all OD pairs.

4.2.5 Price of Satisficing

The price of anarchy (PoA) compares the performances of a satisficing user equilibrium ($C(\mathbf{f}^\kappa)$) against that of a system optimum ($C(\mathbf{f}^*)$). Among possibly multiple satisficing user equilibrium traffic patterns, we are interested in the worst-case. Let $\Psi_\kappa(G, \mathbf{Q}, \mathbf{t})$ be the set of all satisficing user equilibria with a multiplicative factor κ where $G, \mathbf{Q}$, and $\mathbf{t}$ denote the underling network, demand vector, and travel time function, respectively. Then, the PoA for the triplet $(G, \mathbf{Q}, \mathbf{t})$ is defined as follows:

$$\mathsf{PoA}_\kappa(G, \mathbf{Q}, \mathbf{t}) = \max_{\mathbf{f}^\kappa \in \Psi_\kappa(G,\mathbf{Q},\mathbf{t})} \frac{C(\mathbf{f}^\kappa)}{C(\mathbf{f}^*)}, \tag{4.10}$$

where $\mathbf{f}^*$ is the system optimum flow for $(G, \mathbf{Q}, \mathbf{t})$. We are usually interested in its upper bound over a set of triplets, Ω., i.e.

$$\sup_{(G,\mathbf{Q},\mathbf{t})\in\Omega} \mathsf{PoA}_\kappa(G, \mathbf{Q}, \mathbf{t})$$

In the context of bounded rationality and satisficing, we are more interested in comparing the performance of approximate Nash equilibrium $C(\mathbf{f}^\kappa)$ and the performance of the perfectly rational user equilibrium $C(\mathbf{f}^0)$. We define the price of satisifcing (PoSat) of instance $(G, \mathbf{Q}, \mathbf{t})$ as follows:

$$\mathsf{PoSat}_\kappa(G, \mathbf{Q}, \mathbf{t}) = \max_{\mathbf{f}^\kappa \in \Psi_\kappa(G,\mathbf{Q},\mathbf{t})} \frac{C(\mathbf{f}^\kappa)}{C(\mathbf{f}^0)}, \tag{4.11}$$

and its upper bound over Ω is

$$\sup_{(G,\mathbf{Q},\mathbf{t})\in\Omega} \mathsf{PoSat}_{\kappa}(G,\mathbf{Q},\mathbf{t})$$

We define Ω is a set of all triplets where G is a directed graph with a finite number of nodes and arcs, $\mathbf{Q}$ is a vector of finite and positive constants, and $\mathbf{t}(\cdot)$ is a vector of polynomial functions with nonnegative coefficients and of order $n \geq 0$. To emphasize the latter, we also write $\Omega(n)$ instead of Ω when appropriate.

4.3 User Equilibrium with Perception Errors

Related to MSatUE is the user equilibrium with perception error (UE-PE) model. In this model, we assume that network users are optimizing, i.e. seeking the shortest path; however, we assume that users may have their own perception of the travel time function.

We let ϵ_a^w denote the perception error of travel time along arc a of users in OD pair w. A vector $\overline{\mathbf{x}} \in \mathbf{X}$ is a solution to the UE-PE model, if

$$\sum_{a\in\mathcal{A}}\sum_{w\in\mathcal{W}} (t_a(\overline{\mathbf{v}}) - \epsilon_a^w)(x_a^w - \overline{x}_a^w) \geq 0 \qquad \forall \mathbf{x} \in \mathbf{X} \tag{4.12}$$

The above variational inequality assumes that $\boldsymbol{\epsilon}$ is sufficiently small, i.e., $0 \leq \epsilon_a^w < t_a(\overline{v})$ for all $a \in \mathcal{A}, w \in \mathcal{W}$. Under such an assumption, $t_a(\overline{\mathbf{v}}) - \epsilon_a^w$ can be viewed as the *perceived* travel time for arc a for drivers of OD pair w.

The term ϵ_a^w represents the perception error for arc a and OD pair w. In this model, we assume all drivers for each OD pair are homogeneous in their perception of arc travel time.

With changes of variables $\lambda_a^w t_a(\mathbf{v}) = t_a(\mathbf{v}) - \epsilon_a^w$, the UE-PE model (4.12) can be restated as follows:

$$(\mathsf{UE\text{-}PE\text{-}}\mathbf{X}) \qquad \sum_{a\in\mathcal{A}}\sum_{w\in\mathcal{W}} \lambda_a^w t_a(\overline{\mathbf{v}})(x_a^w - \overline{x}_a^w) \geq 0 \qquad \forall \mathbf{x} \in \mathbf{X} \tag{4.13}$$

for some $\boldsymbol{\lambda}$ such that $\lambda_a^w \in (0,1]$ for all $w \in \mathcal{W}$ and $a \in \mathcal{A}$. We observe that the UE-PE model generates a subset of MSatUE traffic flow patterns.

Lemma 4.1 (UE-PE-$\mathbf{X}$ $\implies$ MSatUE)**.** Suppose $\overline{\mathbf{x}}$ is a solution to UE-PE-$\mathbf{X}$ in (4.13) with some $\overline{\boldsymbol{\lambda}}$ where $\overline{\lambda}_a^w \in [\frac{1}{1+\kappa}, 1]$ for all $w \in \mathcal{W}$ and $a \in \mathcal{A}$. Then any $\overline{\mathbf{f}}$ with $\overline{\mathbf{x}} \overset{\text{any}}{\longmapsto} \overline{\mathbf{f}}$ is a κ-MSatUE flow.

We can also provide a path-based formulation of UE-PE:

$$(\textsf{UE-PE-}\mathbf{F}) \quad \sum_{w \in \mathcal{W}} \sum_{p \in \mathcal{P}_w} \widetilde{c}_p^w(\overline{\mathbf{f}})(f_p - \overline{f}_p) \geq 0 \qquad \forall \mathbf{f} \in \mathbf{F} \tag{4.14}$$

for the perceived path travel time functions $\widetilde{c}_p^w(\mathbf{f}) = \sum_{a \in \mathcal{A}} \delta_a^p \lambda_a^w t_a(\mathbf{v})$ with *some* $\boldsymbol{\lambda}$ such that $\lambda_a^w \in (0,1]$ for all $a \in \mathcal{A}, w \in \mathcal{W}$.

Lemma 4.2 (UE-PE-$\mathbf{F}$ $\iff$ UE-PE-$\mathbf{X}$)**.** If $\overline{\mathbf{f}} \in \mathbf{F}$ is a solution to UE-PE-$\mathbf{F}$ in (4.14) for some $\boldsymbol{\lambda}$ such that $\lambda_a^w \in [\frac{1}{1+\kappa}, 1]$, then $\overline{\mathbf{x}}$ with $\overline{\mathbf{f}} \mapsto \overline{\mathbf{x}}$ is a solution to UE-PE-$\mathbf{X}$ in (4.13). Conversely, if $\overline{\mathbf{x}} \in \mathbf{X}$ is a solution to UE-PE-$\mathbf{X}$ in (4.13), then any $\overline{\mathbf{f}}$ with $\overline{\mathbf{x}} \overset{\text{any}}{\longmapsto} \overline{\mathbf{f}}$ is a solution to UE-PE-$\mathbf{F}$ in (4.14).

When the values of λ_a^w are the same across all $w \in \mathcal{W}$, i.e. $\lambda_a = \lambda_a^w$ for all $w \in \mathcal{W}$, we can simplify (4.13) as follows:

$$(\textsf{UE-PE-}\mathbf{V}) \quad \sum_{a \in \mathcal{A}} \lambda_a t_a(\overline{\mathbf{v}})(v_a - \overline{v}_a) \geq 0 \qquad \forall \mathbf{v} \in \mathbf{V} \tag{4.15}$$

for *some* $\boldsymbol{\lambda}$ such that $\lambda_a \in (0,1]$ for each $a \in \mathcal{A}$. The simplified model (4.15) has been considered in the literature for approximate Nash equilibrium (Christodoulou et al., 2011) and Nash equilibrium with deviated travel time functions (Kleer and Schäfer, 2016). For the simplified model, we can state:

Lemma 4.3 (UE-PE-$\mathbf{V}$ $\implies$ UE-PE-$\mathbf{X}$)**.** Suppose that $\overline{v} \in \mathbf{V}$ is a solution to UE-PE-$\mathbf{V}$ in (4.15) for some $\boldsymbol{\lambda}$ such that $\lambda_a \in [\frac{1}{1+\kappa}, 1]$ for all $a \in \mathcal{A}$. Let $\overline{\mathbf{x}}$ be any vector with $\overline{\mathbf{v}} \overset{\text{any}}{\longmapsto} \overline{\mathbf{x}}$. Then $\overline{\mathbf{x}}$ is a solution to UE-PE-$\mathbf{X}$ in (4.13).

$$
\begin{array}{ccccccc}
\textsf{UE-PE-}\mathbf{V} & \Longrightarrow & \textsf{UE-PE-}\mathbf{X} & \Longrightarrow & \text{MSatUE} & \Longrightarrow & (4.16) \\
 & & \Updownarrow & & & & \\
 & & \textsf{UE-PE-}\mathbf{F} & & & &
\end{array}
$$

Figure 4.1 – Summary of lemmas 4.1–4.4.

While Lemmas 4.1, 4.2, and 4.3 provide sufficient conditions for a traffic flow pattern to be a κ-MSatUE, Theorem 1 of Christodoulou et al. (2011) provides a necessary condition. Although Christodoulou et al. (2011) assumed separable arc travel time functions, their proof is still valid for nonseparable travel time functions.

Lemma 4.4 (A necessary condition of MSatUE)**.** Let $\mathbf{f}^\kappa \in \mathbf{F}$ be a κ-MSatUE and $\mathbf{v}^\kappa \in \mathbf{V}$ be the corresponding arc flow vector with $\mathbf{f}^\kappa \mapsto \mathbf{v}^\kappa$. Then we have

$$\sum_{a \in \mathcal{A}} t_a(\mathbf{v}^\kappa)((1+\kappa)v_a - v_a^\kappa) \geq 0 \qquad \forall \mathbf{v} \in \mathbf{V}. \tag{4.16}$$

Christodoulou et al. (2011) derive a tight bound on the price of anarchy on approximate Nash equilibria based on Lemma 4.4.

Figure 4.1 summarizes the relationships between problems addressed in Lemma 4.1 to 4.4. The relation $\mathsf{X} \implies \mathsf{Y}$ means that any solution to X yields a solution to Y Note that this result does not mean that the notions of UE-PE and SatUE are equivalent in general. Rather, the specific modeling of UE-PE-$\mathbf{X}$, UE-PE-$\mathbf{F}$, and UE-PE-$\mathbf{V}$ with the perception error interval $[\frac{1}{1+\kappa}, 1]$, as defined in (4.13), (4.14), and (4.15), respectively, leads to the result in Figure 4.1. If we use a different definition of perception error sets, such a result may not hold.

4.4 Bounding the Price of Satisficing

We first provide analytical bounds of $C(\mathbf{f}^\kappa)$ compared to $C(\mathbf{f}^0)$.

4.4.1 Lessons from the Price of Anarchy

We first observe that $\mathsf{PoSat}_\kappa(G, \mathbf{Q}, \mathbf{t}) \leq \mathsf{PoA}_\kappa(G, \mathbf{Q}, \mathbf{t})$ for any network instance $(G, \mathbf{Q}, \mathbf{t})$, since $C(\mathbf{f}^0) \geq C(\mathbf{f}^*)$. This enables us to use the results from the price of anarchy literature for bounding PoSat. Theorem 2 of Christodoulou et al. (2011) bound the price of anarchy when arc travel-time functions are separable and polynomial with nonnegative coefficients and of degree n that leads immediately to the following result:

Lemma 4.5. Suppose $\mathbf{f}^\kappa$ is a κ-MSatUE flow, and $t_a(\cdot)$ is polynomial with nonnegative coefficients and of degree n. Define

$$\zeta(\kappa, n) = \begin{cases} (1+\kappa)^{n+1} & \text{if } \kappa \geq (n+1)^{1/n} - 1, \\ \left(\frac{1}{1+\kappa} - \frac{n}{(n+1)^{(n+1)/n}}\right)^{-1} & \text{if } 0 \leq \kappa \leq (n+1)^{1/n} - 1. \end{cases} \tag{4.17}$$

Then we have

$$C(\mathbf{f}^*) \leq C(\mathbf{f}^\kappa) \leq \zeta(\kappa, n) C(\mathbf{f}^*) \leq \zeta(\kappa, n) C(\mathbf{f}^0). \tag{4.18}$$

That is, the PoSat is bounded above by $\zeta(\kappa, n)$.

The bound in Lemma 4.5 is not tight when κ is small. For example, when $\kappa = 0, C(\mathbf{f}^{\boldsymbol{\kappa}}) = C(\mathbf{f}^{\mathbf{0}})$. Thus, $\frac{C(\mathbf{f}^{\boldsymbol{\kappa}})}{C(\mathbf{f}^{\mathbf{0}})} = 1$. However, (19) yields the following:

$$\frac{C(\mathbf{f}^{\boldsymbol{\kappa}})}{C(\mathbf{f}^{\mathbf{0}})} \leq \left(1 - \frac{n}{(n+1)^{\frac{n+1}{n}}}\right)^{-1}$$

where the expression on the right is strictly larger than one. For example, the expression reduces to $\frac{4}{3}$ when $n = 1$ and approaches infinity when n is large.

In Lemma 3 of Christodoulou et al. (2011), the existence of a network instance with $C(\mathbf{f}^\kappa) = (1+\kappa)^{n+1} C(\mathbf{f}^*)$ is shown for $\kappa \geq (n+1)^{1/n} - 1$ via a circular network example presented in Figure 4.2. The circular network includes $m + l$ nodes where positive integers m and l are chosen so that $\frac{m}{l} = 1 + \kappa$. All nodes lie in a circle and each node i is adjacent to two neighboring nodes via arc $(i, i+1)$ and $(i-1, i)$. The arc cost function for arc a is

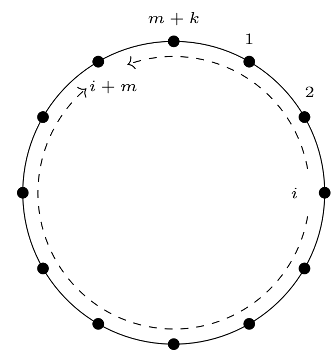


Figure 4.2 – Circular network of Christodoulou et al. (2011)

$t_a(v_a) = (v_a)^n$, where v_a is the total arc flow in arc a. There are $m+l$ OD pairs $(i, i+m)$ for $i = 1, 2, .., m+l$, with unit demand from node i to node $i+m$ (indices are taken cyclically). Note that the circular network can be easily converted to a directed network by replacing each undirected arc with two directed arcs with opposite direction. The cost associated with arc a will be $t_a(v_a, \widehat{v}_a) = (v_a + \widehat{v}_a)^n$ in this case where $\widehat{v}_a$ is the flow in the arc with opposite direction.

For each OD pair $(i, i+m)$, there are two paths, clockwise and counterclockwise. The former contains m arcs and the latter has only l. Note that our choice requires that $\frac{m}{l} = (1+\kappa)$ where $\kappa \geq 0$, i.e., $m \geq l$. Consider the all-or-nothing strategy that sends the unit demand for every OD pair along only one path, clockwise or counterclockwise. Using the clockwise strategy, each arc has m units of flows and costs m^n. Thus, the clockwise path costs $m \cdot m^n$, while the cost of the counterclockwise one is $l \cdot m^n$. Because $\frac{m \cdot m^n}{l \cdot m^n} = \frac{m}{l} \geq 1$, the clockwise strategy is not in a user equilibrium unless $m = l$. For the counterclockwise strategy, each arc has l units of flows instead and costs l^n. Similarly, the clockwise and counterclockwise path cost $m \cdot l^n$ and $l \cdot l^n$, respectively. Using the same reasoning as before, flow-bearing (or the counterclockwise) paths are less expensive than or the same as the unused ones. Thus, the counterclockwise strategy is in user equilibrium and yields $(m+l) \cdot l^{n+1}$ as the total travel cost. Consider MSatUE. Because it is PRUE, the counterclockwise strategy is automatically in MSatUE. But, the clockwise one is also in MSatUE because the cost of flow-bearing (or the clockwise) path is exactly $(1+\kappa)$ time the cost of the shortest path, i.e., the counterclockwise one. Additionally, the total travel cost of the clockwise strategy is $(m+l) \cdot m^{n+1}$. Then, the PoSat of this circular network is

$$\frac{(m+l) \cdot m^{n+1}}{(m+l) \cdot l^{n+1}} = \left(\frac{m}{l}\right)^{n+1} = (1+\kappa)^{n+1}.$$

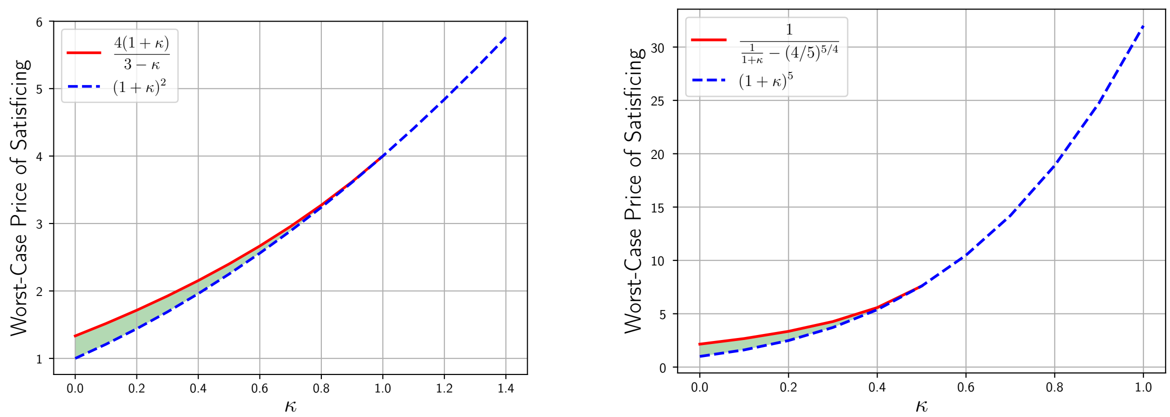
Thus, the bound is tight. Now consider the system problem for the circular network. The problem objective is to minimize $\sum_{a \in \mathcal{A}} v_a t_a(\mathbf{v})$. Using the counterclockwise strategy, the partial derivative of the objective function with respect to v_a is $(n+1)l^n$. Then, switching to the clockwise path would increase the total travel cost by $(m-l)(n+1)l^n \geq 0$ per unit flow. Thus, the counterclockwise strategy is system optimal because switching to the unused path does not result in a reduction in the total travel time. Observe that the $\textsf{UE-PE-}\mathbf{X}$ model can capture the satisficing behavior of network users in the circular network adequately. Let us consider $\boldsymbol{\lambda}$ as follow as (indices are taken cyclically):

$$\lambda_a^{w_i} = \begin{cases} \frac{1}{1+\kappa} & \text{for } a \in \{(j, j+1) : j = i, i+1, ..., i+m-1\} \\ 1 & \text{for } a \in \{(j-1, j) : j = i, i-1, ..., i-l+1\} \end{cases}$$

Under the above $\boldsymbol{\lambda}$, if all network users choose the clockwise path, the flow in each arc will be equal to m, the path cost for each OD pair will be $\frac{1}{1+k} m^{n+1} = lm^n$, which is equal to the cost of the alternative path, and thus the clockwise path is a solution to $\textsf{UE-PE-}\mathbf{X}$ and $\textsf{PoSat}_\kappa = (1+\kappa)^{n+1}$. In Section 4.5, we will compute the PoSat numerically for these examples to confirm that $\textsf{UE-PE-}\mathbf{X}$ is a useful model to find $\textsf{PoSat}_\kappa$.

By Lemma 4.5, when travel time functions are polynomials of degree n with nonnegative coefficients, the PoSat is bounded as follows:

$$\textsf{PoSat}_\kappa(G, \mathbf{Q}, \mathbf{t}) \leq \begin{cases} (1+\kappa)^{n+1} & \text{if } \kappa \geq (n+1)^{1/n} - 1, \\ \left(\frac{1}{1+\kappa} - \frac{n}{(n+1)^{(n+1)/n}}\right)^{-1} & \text{if } 0 \leq \kappa \leq (n+1)^{1/n} - 1. \end{cases} \tag{4.19}$$



(a) $n = 1$

(b) $n = 4$

Figure 4.3 – The worst-case price of satisficing for $n = 1$ and $n = 4$.

for all $(G, \mathbf{Q}, \mathbf{t}) \in \Omega(n)$ and by the circular network example in Figure 4.2, we know that there indeed exists a network instance $(G, \mathbf{Q}, \mathbf{t}) \in \Omega(n)$ such that $\mathsf{PoSat}_\kappa(G, \mathbf{Q}, \mathbf{t}) = (1 + \kappa)^{n+1}$ for all $\kappa \geq 0$. Therefore when $\kappa \geq (n + 1)^{1/n} - 1$, the bound in (4.19) is tight. Figure 4.3a shows the bounds in (4.19) when travel time functions are linear or when $n = 1$. For smaller κ values, the worst-case PoSat falls in the shaded interval, while for larger κ values, it is exactly $(1 + \kappa)^2$. Figure 4.3b shows the same bounds when $n = 4$ instead. When κ is zero, we have $\mathbf{f}^\kappa = \mathbf{f}^0$; hence, we must have the PoSat approach to 1. With this observation, we naturally ask a question: Does $(1 + \kappa)^{n+1}$ provide a tight bound on PoSat_κ for all $\kappa \geq 0$? We present partial answers to this question in the following sections.

4.4.2 Increased Travel Demands and Travel Time Functions

We first define new sets of flow vectors. When the travel demand Q_w for each $w \in \mathcal{W}$ is multiplied by the factor $1 + \kappa$, we define

$$\mathbf{F}_{1+\kappa} = \left\{ \mathbf{f} : \sum_{p \in \mathcal{P}_w} f_p = (1 + \kappa) Q_w \quad \forall w \in \mathcal{W}, \qquad f_p \geq 0 \quad \forall p \in \mathcal{P} \right\},$$

$$\mathbf{V}_{1+\kappa} = \left\{ \mathbf{v} : v_a = \sum_{p \in \mathcal{P}} \delta_a^p f_p \quad \forall a \in \mathcal{A}, \qquad \mathbf{f} \in \mathbf{F}_{1+\kappa} \right\},$$

$$\mathbf{X}_{1+\kappa} = \left\{ \mathbf{x} : x_a^w = \sum_{p \in \mathcal{P}_w} \delta_a^p f_p \quad \forall a \in \mathcal{A}, w \in \mathcal{W} \qquad \mathbf{f} \in \mathbf{F}_{1+\kappa} \right\}.$$

The above three sets can equivalently be written as follows:

$$\mathbf{F}_{1+\kappa} = \{(1+\kappa)\mathbf{f} : \mathbf{f} \in \mathbf{F}\},$$
$$\mathbf{V}_{1+\kappa} = \{(1+\kappa)\mathbf{v} : \mathbf{v} \in \mathbf{V}\},$$
$$\mathbf{X}_{1+\kappa} = \{(1+\kappa)\mathbf{x} : \mathbf{x} \in \mathbf{X}\}.$$

We will use 'hat' for flow vectors in these sets, for example, $\widehat{\mathbf{f}}^{\kappa} \in \mathbf{F}_{1+\kappa}$, while without hat in the original sets as in $\mathbf{f}^{\kappa} \in \mathbf{F}$.

We consider cases when the travel time functions $t_a(\cdot)$ are polynomials of order n, in particular, the following form of *asymmetric* arc travel time function for each $a \in \mathcal{A}$:

$$\begin{aligned} t_a(\mathbf{v}) &= \sum_{m=0}^{n} b_{am} \left(\sum_{e \in \mathcal{A}} d_{aem} v_e \right)^m \\ &= \sum_{m=0}^{n} b_{am} \left(\mathbf{d}_{am}^{\top} \mathbf{v} \right)^m \end{aligned} \tag{4.20}$$

for some constants b_{am} for $m = 0, 1, ..., n$ and d_{aem} for $e \in \mathcal{A}$ and $m = 0, 1, ..., n$. Note that we use the vector form $\mathbf{d}_{am} = (d_{aem} : e \in \mathcal{A})$. The travel time function (4.20) is a general form of the travel time functions considered in the traffic equilibrium literature (Meng et al., 2014; Panicucci et al., 2007). If $\mathbf{d}_{am}$ is a unit vector such that d_{aem} is 1 if $a = e$ and 0 otherwise, we have a separable polynomial arc travel time function that has been used in the literature popularly (Christodoulou et al., 2011; Roughgarden and Tardos, 2002):

$$t_a(v_a) = \sum_{m=0}^{n} b_{am}(v_a)^m = b_{a0} + b_{a1}v_a + b_{a2}(v_a)^2 + \cdots + b_{an}(v_a)^n. \tag{4.21}$$

Lemma 4.6. With the polynomial travel time function (4.20), for any $\mathbf{f} \in \mathbf{F}$, we have

$$C((1+\kappa)\mathbf{f}) \leq (1+\kappa)^{n+1} C(\mathbf{f}) \tag{4.22}$$

for all $\kappa \geq 0$ and $n \geq 0$.

4.4.3 Cases with Separable, Monomial Arc Travel Time Functions

As a simple case, we consider separable, monomial functions of degree n for arc travel time of the following form:

$$t_a(v_a) = b_a (v_a)^n \tag{4.23}$$

with a positive scalar b_a for each $a \in \mathcal{A}$ and nonnegative constant n.

It is well known (Beckmann et al., 1956) that $\mathbf{v}^0 \in \mathbf{F}$ is a user equilibrium flow, if and only if it minimizes the following potential function

$$\Phi(\mathbf{v}) = \sum_{a \in \mathcal{A}} \int_0^{v_a} t_a(u) \, \mathrm{d}u = \sum_{a \in \mathcal{A}} \frac{b_a}{n+1} (v_a)^{n+1}$$

when the arc travel time functions are separable, so that the integral is well defined. Similarly, $\mathbf{v}^\kappa \in \mathbf{V}$ is a κ-MSatUE flow, if it is a solution to $\mathsf{UE\text{-}PE\text{-}}\mathbf{V}$, or equivalently, if it minimizes the following potential function (Christodoulou et al., 2011)

$$\Psi(\mathbf{v}; \boldsymbol{\lambda}) = \sum_{a \in \mathcal{A}} \int_0^{v_a} \lambda_a t_a(u) \, \mathrm{d}u = \sum_{a \in \mathcal{A}} \frac{\lambda_a b_a}{n+1} (v_a)^{n+1}$$

for some $\lambda_a \in [\frac{1}{1+\kappa}, 1]$ for each $a \in \mathcal{A}$.

When travel time functions are separable, we can show the following result (Englert et al., 2010; Takalloo and Kwon, 2018):

Lemma 4.7. When the arc travel time functions are in the form of (4.21), let $\mathbf{f}^0 \in \mathbf{F}$ and $\widehat{\mathbf{f}}^0 \in \mathbf{F}_{1+\kappa}$ be the PRUE flows with the corresponding travel demands. We can show

$$C(\widehat{\mathbf{f}}^0) \leq (1+\kappa)^{n+1} C(\mathbf{f}^0) \tag{4.24}$$

for all $\kappa \geq 0$ and $n \geq 0$.

Although Englert et al. (2010) consider cases with a single OD pair only with interest in the changes in the path travel time, the same technique can be used to prove Lemma 4.7 for cases with multiple OD pairs.

Using Lemma 4.7, we show that a solution to UE-PE-**V** is an MSatUE flow.

Theorem 4.1. When the arc travel time functions are of the form (4.23), let $\overline{\mathbf{v}} \in \mathbf{V}$ be a solution to UE-PE-**V** and $\overline{\mathbf{f}} \in \mathbf{F}$ is the any corresponding path flow with $\overline{\mathbf{v}} \overset{\text{any}}{\longmapsto} \overline{\mathbf{f}}$. We let $\widehat{\mathbf{f}}^0 \in \mathbf{F}_{1+\kappa}$ be the PRUE flows. Then we have $C(\mathbf{f}^\kappa) \leq C(\widehat{\mathbf{f}}^0)$, and consequently $C(\mathbf{f}^\kappa) \leq (1+\kappa)^{n+1} C(\mathbf{f}^0)$ for all $\kappa \geq 0$.

Note that the bound obtained in Theorem 4.1 relies on the sufficient condition, not a necessary condition. Therefore, the result is not applicable to all MSatUE flows, although it provides a useful bound in the framework of UE-PE models.

4.4.4 Cases with Separable Arc Travel Time Functions

We consider general polynomial, separable arc travel functions in the form of (4.21).

Theorem 4.2. Suppose that the arc travel time functions are in the form of (4.21). Let $\mathbf{f}^\kappa \in \mathbf{F}$ be any κ-MSatUE and $\widehat{\mathbf{f}}^0 \in \mathbf{F}_{1+\kappa}$ be the PRUE flow. Suppose that $\kappa \geq 0$ is sufficiently small, in particular, so that

$$\sum_{p \in \mathcal{P}} [c_p(\widehat{\mathbf{f}}^0) - c_p(\mathbf{f}^\kappa)](\widehat{f}_p^0 - f_p^\kappa) \geq \kappa \sum_{p \in \mathcal{P}} c_p(\mathbf{f}^\kappa) \left| \widehat{f}_p^0 - f_p^\kappa \right|. \tag{4.25}$$

Then we have $C(\mathbf{f}^\kappa) \le C(\widehat{\mathbf{f}}^0)$. Consequently $C(\mathbf{f}^\kappa) \le (1+\kappa)^{n+1} C(\mathbf{f}^0)$, and

$$\sup_{(G,\mathbf{Q},\mathbf{t})\in\Omega(n)} \mathsf{PoSat}_\kappa(G,\mathbf{Q},\mathbf{t}) = (1+\kappa)^{n+1}.$$

Theorem 4.2 depends on condition (4.25) and a similar condition appears in general asymmetric cases as in Theorem 4.3. We discuss this condition in Section 4.4.6.

4.4.5 General Cases with Asymmetric Arc Travel Time Functions

We consider asymmetric arc travel time functions (4.20), in which case Lemma 4.7 is not applicable. We first observe that the multiple of a PRUE flow, $(1+\kappa)\mathbf{f}^0$, provides a satisficing solution to the traffic equilibrium problem with the increased travel demand.

Lemma 4.8. Suppose $t_a(\cdot)$ are polynomials of order n as defined (4.20). If $\mathbf{f}^0 \in \mathbf{F}$ is a PRUE flow, then $(1+\kappa)\mathbf{f}^0$ is a σ-MSatUE flow with $\sigma = (1+\kappa)^n - 1$ in $\mathbf{F}_{1+\kappa}$. When $n = 1$, we have $\sigma = \kappa$.

By introducing an additional condition, we compare MSatUE flows with the proportional travel demand increase, and obtain the worst-case bound of PoSat.

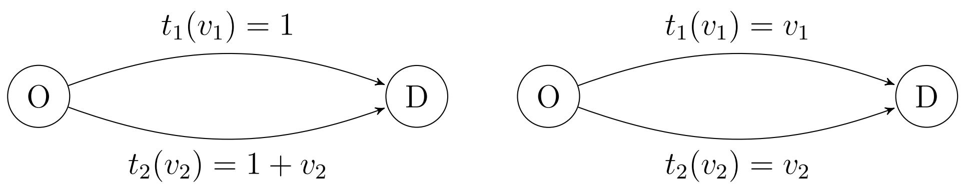
Theorem 4.3. Let $\mathbf{f}^\kappa \in \mathbf{F}$ be any κ-MSatUE and $\widehat{\mathbf{f}}^\sigma \in \mathbf{F}_{1+\kappa}$ be any σ-MSatUE flows with the corresponding travel demands, when $\sigma = (1+\kappa)^n - 1$. Suppose that $\kappa \ge 0$ is sufficiently small, in particular, so that

$$\sum_{p\in\mathcal{P}} [c_p(\widehat{\mathbf{f}}^\sigma) - c_p(\mathbf{f}^\kappa)](\widehat{f}_p^\sigma - f_p^\kappa) \ge \sigma \sum_{p\in\mathcal{P}} \max\{c_p(\widehat{\mathbf{f}}^\sigma), c_p(\mathbf{f}^\kappa)\} \left| \widehat{f}_p^\sigma - f_p^\kappa \right|. \qquad (4.26)$$

Then we have $C(\mathbf{f}^\kappa) \le C(\widehat{\mathbf{f}}^\sigma)$. Consequently $C(\mathbf{f}^\kappa) \le (1+\kappa)^{n+1} C(\mathbf{f}^0)$, and

$$\sup_{(G,\mathbf{Q},\mathbf{t})\in\Omega(n)} \mathsf{PoSat}_\kappa(G,\mathbf{Q},\mathbf{t}) = (1+\kappa)^{n+1}.$$

Note that condition (4.26) is stronger than condition (4.25) for separable travel time functions. This is natural, since we consider more general classes of travel time functions.



(a) Example 1 (b) Example 2

Figure 4.4 – Simple illustrative examples for evaluating PoSat where the travel demand is Q from node O to node D.

4.4.6 Illustrative Examples

For the illustration purpose, we consider two examples in Figure 4.4 with linear travel time functions, where $n = 1$. In Example 1, the travel time function in the first arc is not increasing. We can verify that

$$\max C(\mathbf{f}^\kappa) = \begin{cases} Q + \kappa^2 & \text{if } \kappa \leq Q, \qquad \text{with } \mathbf{f}^\kappa = (Q - \kappa, \kappa) \\ (1+Q)Q & \text{if } \kappa \geq Q, \qquad \text{with } \mathbf{f}^\kappa = (0, Q) \end{cases}$$

among all κ-MSatUE flows in $\mathbf{F}$ and

$$C(\widehat{\mathbf{f}}^0) = (1+\kappa)Q \qquad \text{with } \widehat{\mathbf{f}}^0 = (1+\kappa)\mathbf{f}^0 = ((1+\kappa)Q, 0).$$

among all κ-MSatUE flows in $\mathbf{F}_{1+\kappa}$. Comparing the two quantities, we observe $C(\mathbf{f}^\kappa) \leq C(\widehat{\mathbf{f}}^0)$ in both cases. To prove Theorem 4.3, condition (4.25) needs to hold only for these two flow vectors. Regardless of the value of κ, however, it is impossible to satisfy condition (4.25), although the worst-case PoSat bound $(1+\kappa)^{n+1}$ still holds for all $\kappa \geq 0$. The price of satisficing is $1 + \frac{\kappa^2}{Q}$ if $\kappa < Q$ and $1 + Q$ if $\kappa \geq Q$ in this example, both of which are less than $(1+\kappa)^2$.

On the other hand, in Example 2, we have strictly monotone travel time functions in both arcs. Similarly, we consider

$$\max C(\mathbf{v}^\kappa) = \frac{2+2\kappa+\kappa^2}{(2+\kappa)^2}Q \qquad \text{with } \mathbf{f}^\kappa = \left(\frac{Q}{2+\kappa}, \frac{(1+\kappa)Q}{2+\kappa}\right)$$
$$C(\widehat{\mathbf{v}}^0) = \frac{(1+\kappa)^2}{2}Q \qquad \text{with } \widehat{\mathbf{f}}^0 = (1+\kappa)\mathbf{f}^0 = \left(\frac{(1+\kappa)Q}{2}, \frac{(1+\kappa)Q}{2}\right)$$

and can verify that $C(\mathbf{f}^\kappa) \leq C(\widehat{\mathbf{f}}^0)$ for all $\kappa \geq 0$. In Example 2, we note that (4.26) holds for $\kappa \leq 0.206$. In this example, we observe that the price of satisficing is $\frac{2(2+2\kappa+\kappa^2)}{(2+\kappa)^2}$, which is no greater than $(1+\kappa)^2$ for all $\kappa \geq 0$.

4.4.7 Other Approaches

When there is a single origin and multiple destinations, i.e., a single common origin node, in the network, Kleer and Schäfer (2016) introduces the notion of the *deviation ratio* that compares the system performances of the user equilibrium and the equilibrium with *deviated* travel time functions $\widetilde{t}_a(\cdot)$. The notion of deviation may also be interpreted as perception in our definition. In a special case, the deviation ratio is reduced to the price of risk aversion (Nikolova and Stier-Moses, 2015) that compares the performances of equilibria among risk-averse and risk-neutral network users.

Kleer and Schäfer (2016) define the (separable) deviated travel time functions with the following bounds:

$$t_a(v_a) + \alpha t_a(v_a) \leq \widetilde{t}_a(v_a) \leq t_a(v_a) + \beta t_a(v_a) \tag{4.27}$$

where $-1 \leq \alpha \leq 0 \leq \beta$. The consideration of this deviated travel time function generalizes our UE-PE model where $\alpha = -\frac{\kappa}{1+\kappa}$ and $\beta = 0$. Kleer and Schäfer (2016) show that the worst-case deviation ratio with (4.27) is bounded by

$$1 + \frac{\beta - \alpha}{1+\alpha}\left\lceil \frac{|\mathcal{N}| - 1}{2} \right\rceil Q.$$

Therefore, we obtain the following theorem:

Theorem 4.4 (Kleer and Schäfer, 2016)**.** Consider a directed graph with a single common origin node with the total travel demand Q and let $|\mathcal{N}|$ be the number of nodes. Then we have

$$\frac{Z(\mathbf{v}^{\kappa})}{Z(\mathbf{v}^{0})} \leq 1 + \kappa \left\lceil \frac{|\mathcal{N}| - 1}{2} \right\rceil Q \tag{4.28}$$

where $\mathbf{v}^{\kappa}$ is a solution to UE-PE-**V** in (4.15).

Note that Theorem 4.4 only covers a subset of the entire MSatUE flows, as it is limited to the solutions UE-PE-**V** in (4.15) and is applicable to cases with a *single* common origin. When Theorem 4.4 is applied in the examples in Figure 4.4, the bound (4.28) becomes $1+\kappa Q$.

4.5 Numerical Bounds

To quantify PoSat in typical traffic networks and compare it with the analytical bound obtained in Theorem 4.3, we define the worst-case problem for the total travel time under MSatUE as follows:

$$\max_{\mathbf{v}^{\kappa}} \quad Z(\mathbf{v}^{\kappa}) = \sum_{a \in \mathcal{A}} z_a(v_a^{\kappa}) = \sum_{a \in \mathcal{A}} t_a(\mathbf{v}^{\kappa}) v_a^{\kappa} \tag{4.29}$$

$$\text{subject to} \quad \mathbf{v}^{\kappa} \text{ is an } \underline{\text{MSatUE}} \text{ flow with } \kappa$$

To quantify the benefit of satisficing, instead of maximizing, we can minimize the objective function (4.29). Since MSatUE involves path-based definition and formulation, (4.29) is numerically more challenging to solve. Instead, we replace MSatUE by UE-PE-**X**. We know that the UE-PE-**X**. models provide a subset of MSatUE traffic flow patterns as seen in Lemmma 4.1; hence by using UE-PE-**X** models, we will obtain suboptimal solutions to (4.29).

Using UE-PE-**X** in (4.12), we formulate the worst-case problem as follows:

$$\max_{\overline{\mathbf{v}},\overline{\mathbf{x}},\boldsymbol{\epsilon}} \quad Z(\overline{\mathbf{v}}) = \sum_{a\in\mathcal{A}} z_a(\overline{\mathbf{v}}) = \sum_{a\in\mathcal{A}} t_a(\overline{\mathbf{v}})\overline{v}_a \tag{4.30}$$

$$\text{subject to} \quad \sum_{a\in\mathcal{A}}\sum_{w\in\mathcal{W}} (t_a(\overline{\mathbf{v}}) - \epsilon_a^w)(x_a^w - \overline{x}_a^w) \geq 0 \qquad \forall \mathbf{x}\in\mathbf{X} \tag{4.31}$$

$$\overline{v}_a^\kappa = \sum_{w\in\mathcal{W}} \overline{x}_a^w \qquad \forall a\in\mathcal{A} \tag{4.32}$$

$$\overline{x}\in\mathbf{X} \tag{4.33}$$

$$0 \leq \epsilon_a^w \leq \frac{\kappa}{1+\kappa} t_a(\overline{\mathbf{v}}) \qquad \forall a\in\mathcal{A} \tag{4.34}$$

Problem (4.30) is an optimization problem with equilibrium constraints. We can replace the equilibrium condition (4.31) by the following KKT conditions to create a single-level optimization problem:

$$t_a(\overline{\mathbf{v}}) - \epsilon_a^w + \pi_i^w - \pi_j^w \geq 0 \qquad \forall w\in\mathcal{W}, a\in\mathcal{A} \tag{4.35}$$

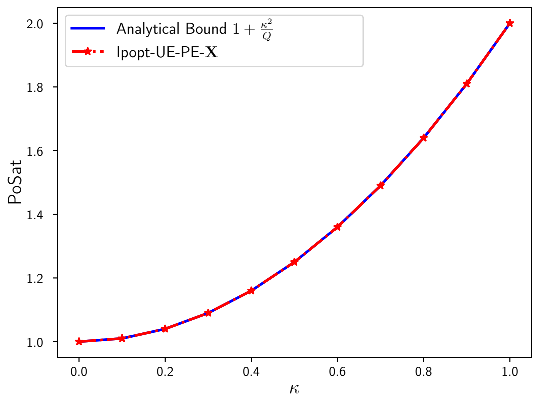
$$\overline{x}_a^w(t_a(\overline{\mathbf{v}}) - \epsilon_a^w + \pi_i^w - \pi_j^w) = 0 \qquad \forall w\in\mathcal{W}, a\in\mathcal{A} \tag{4.36}$$

$$\sum_{a\in\mathcal{A}_i^+} \overline{x}_a^w - \sum_{a\in\mathcal{A}_i^-} \overline{x}_a^w = q_i^w \qquad \forall w\in\mathcal{W}, i\in\mathcal{N} \tag{4.37}$$

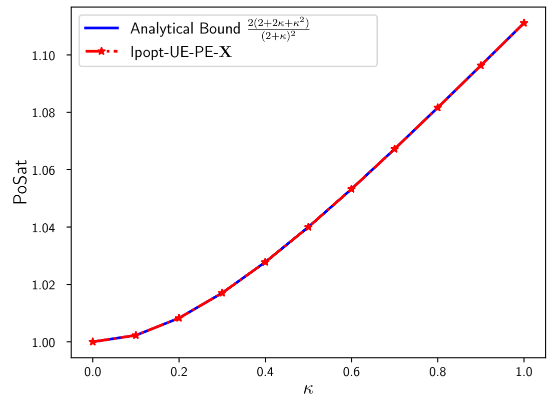
The resulting problem is a mathematical program with complementarity conditions (MPCC), which is nonlinear and nonconvex. Finding a global solution to MPCC problems is in general difficult, and Kleer and Schäfer (2016) has shown that solving the above MPCC optimally is NP-hard. In order to solve this problem, we use an interior point method by utilizing the Ipopt nonlinear solver (Wächter and Biegler, 2006) with multiple starting solutions.

4.5.1 Numerical Experiments

In this section we present some examples to compare the total travel times in MSatUE and PRUE numerically for both separable and asymmetric networks. We approximate MSatUE by UE-PE-**X** and solve it by the Ipopt nonlinear solver, after reformulating (4.29) as a single-



(a) Example 1 with $Q = 1$



(b) Example 2

Figure 4.5 – PoSat for the simple networks in Figure 4.4

level optimization problem using KKT conditions. We use the JuMP package (Dunning et al., 2017b) in Julia Programming Language for modeling and interfacing with the Ipopt solver.

4.5.1.1 Simple Networks

To test the validity and the strength of UE-PE-**X** model, we first consider Examples 1 and 2 in Figure 4.4. Figure 4.5 compares the PoSat under UE-PE-**X** with the PoSat under MSatUE, obtained in Section 4.4.6. As Figure 4.5 shows, the PoSat under UE-PE-**X** is equal to the the PoSat under MSatUE for both examples, which suggests that the UE-PE-**X** model is an effective model.

4.5.1.2 Circular Network of Christodoulou et al. (2011)

We also compute the PoSat under UE-PE-**X** model for the circular network of Christodoulou et al. (2011) presented in Figure 4.2. For numerical experiments, we assign m and l to the smallest positive integers such that $\frac{m}{l} = (1+\kappa)^5$, and $\kappa \in \{0, 0.1, 0.2, ..., 1\}$. We also set $n = 4$. As it can be seen in Figure 4.6, we obtained identical results for circular network under UE-PE-**X**, using the Ipopt solver, which shows that the UE-PE-**X** model can obtain the upper bound provided in Lemma 4.5.

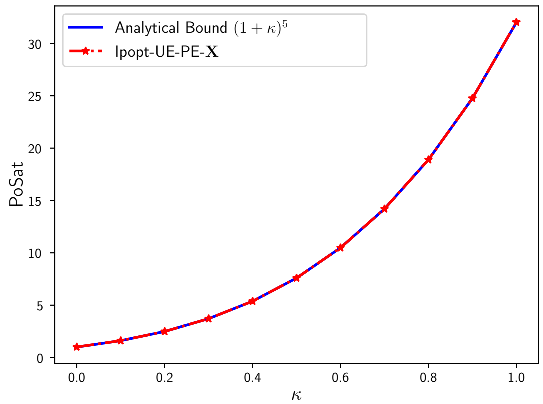
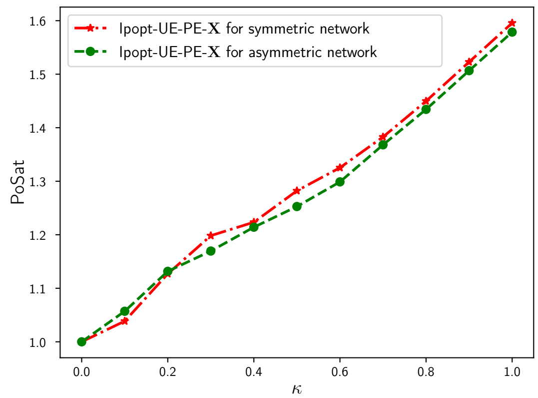


Figure 4.6 – PoSat for the circular network in Figure 4.2

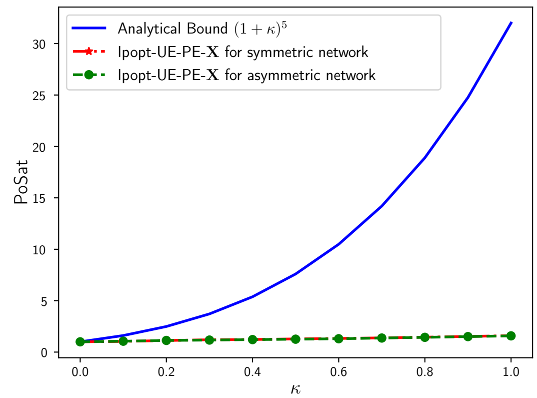
4.5.1.3 Larger Networks

We present some examples to compare the total travel times in MSatUE and PRUE numerically and compare the numerical worst-cases with the analytical bound given in Theorem 4.2 for larger networks with both separable and non-separable, asymmetric arc cost functions. As (4.29) is a non-convex problem, the Ipopt solver can produce a local minimum at best. To obtain a higher-quality local minimum, we solve the problem multiple times by using different initial solution. For generating different initial solutions for the network with separable arc cost functions, we utilize UE-PE-$\mathbf{V}$ model. We generate initial $\boldsymbol{\lambda}$ randomly and use the Frank-Wolfe algorithm to obtain the corresponding $\mathbf{v}$ and $\mathbf{x}$. For the network with non-separable cost function, we can use the fixed point method (Dafermos, 1980) with a randomized $\boldsymbol{\lambda}$ to obtain an initial solution $\mathbf{x}$. We randomly generate five initial starting points for each example and report the largest PoSat values.

We first consider the nine-node network presented in Hearn and Ramana (1998). The nine-node network consists of 9 nodes and 18 arcs, and the travel time functions are polynomials of order $n = 4$. We also create an asymmetric variant of the nine-node network as shown in Figure A.1 in Appendix E. The asymmetric nine-node network has non-separable arc cost function in the form of (A29). The comparison result is presented in Figure 4.7a.



(a) Approximate PoSat (Ipopt-UE-PE-**X**)



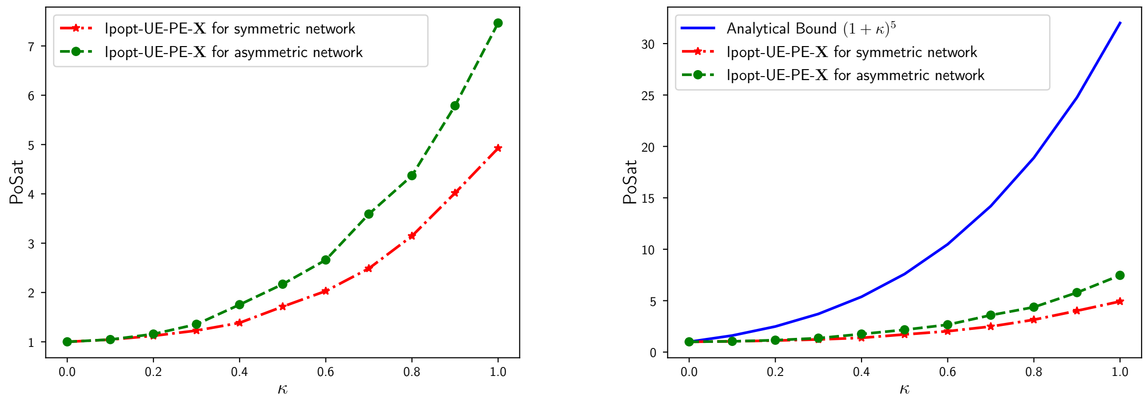
(b) Comparing analytical and numerical PoSat

Figure 4.7 – PoSat for nine-node network

As Figure 4.7a represents PoSat_κ increases with κ for both symmetric and asymmetric nine-node network since PRUE total travel time is fixed with respect to κ, while the worst-case MSatUE total travel time increases as κ increases. Moreover, PoSat_κ is smaller for symmetric nine-node network compared to the asymmetric nine-node network for smaller κ values (0.1 and 0.2), but it is greater for larger κ values ($\kappa \geq 0.3$). In general, the gap between PoSat_κ for symmetric nine-node network and asymmetric nine-node network is small.

Figure 4.7b compares the numerical PoSat_κ with the analytical bound provided in Theorem 4.2 for MSatUE for the nine-node network. We observe that there is a large gap between the analytical and numerical bounds which increases with κ. Although the analytical result certainly provides a valid bound, it is too large to be practically useful in realistic road networks. This indicates opportunities for empirical studies on the analytical bounds that depend on more network-specific information such as travel demands and travel time functions. The bound $(1+\kappa)^{n+1}$ in Theorem 4.2 is independent from such network-specific information.

We also consider the Sioux Falls network presented in Suwansirikul et al. (1987), which consists of 24 nodes, 76 arcs, and 576 OD pairs. The arc travel cost function is the BPR function, which is a polynomial function with degree $n = 4$. We also consider an asymmetric



(a) Approximate PoSat (Ipopt-UE-PE-**X**) (b) Comparing analytical and numerical PoSat

Figure 4.8 – PoSat for Sioux Falls network

variant of Sioux Falls network with arc cost function in the form of (A29). As Figure 4.8a represents, PoSat_κ increases with κ for both symmetric and asymmetric Sioux Falls network, and it is greater compared to the nine-node network for both symmetric and asymmetric networks. Furthermore, PoSat_κ is greater for asymmetric Sioux Falls network compared with the symmetric Sioux Falls network for all positive κ values, and the gap between PoSat_κ for symmetric Sioux Falls network and PoSat_κ for asymmetric Sioux Falls network increases with κ. Figure 4.8b compares the numerical PoSat_κ with the analytical bound. The gap between the analytical and the numerical bound is tighter compared to the nine-node network, but it is still considerable.

Chapter 5: Conclusions and Future Research Directions

In this dissertation, we use game theory approach to solve three separate transportation problems. In the first problem, we design a new auction market for fractional vehicle ownership market that considers spatial information of the bidders. We propose the use of bidder-defined items, which enables bidders to bid on any time intervals they wish. We use clique constraints to formulate the WDP, which enables us to generate the model and solve it faster. For solving the problem for larger instances, we design a greedy algorithm to find a primal solution and provide a maximal-clique relaxation of the problem to obtain a high-quality dual bound for that. We compare the performance of the proposed formulations and solution approaches through extensive numerical experiments. We can extend this work in several directions. First, the methodology of this paper can be applied to other combinatorial auctions, including resource allocation in cloud computing and ride-sharing markets with some variations, etc. Second, We can improve the solution approaches by designing algorithms based on maximal-clique based relaxation. Third, we can improve the vehicle co-ownership auction market by considering other features. For instance, we can consider the uncertainty in the departure and arrival time of the bidders in the auction market. Fourth, as the VCG mechanism does not perform well, especially regarding the auctioneer revenue, one research direction is to design a practical iterative auction for fractional car ownership, which helps preference elicitation and generating higher revenue.

In the second problem, we considered a pessimistic approach for hazmat toll pricing problem. We studied the existence of the solution for pessimistic and approximate pessimistic hazmat toll pricing problems. We developed solution approaches, which are closely related to Wiesemann et al. (2013) and utilize disjunctive programming, to solve the approximate pessimistic hazmat toll pricing problem. To model the disjunctive constraints, we utilize both the Balas formulation and the big-M formulation approaches, both of which use big-M constraints to linearize the bilinear terms. While Balas formulation is ideal and stronger, the practical performance was worse than the big-M formulation with the bounds on big-M constants we developed. Through numerical experiments in a realistic network, we confirmed that there usually exist multiple optimal solutions in the problem, hence requiring a pessimistic approach for robustness. We can further explore this problem in several directions. First, the pessimistic approach can be applied to other bi-level optimization problems including network design hazmat problem and general network pricing problems. Second, it may be possible to improve the bound on M^3_{sij}. Improving the bound on M^3_{sij} makes the master problem easier to solve and possibly can improve the quality of the solution. Moreover, it may be possible to propose new valid inequalities to solve the master problem more efficiently.

In the third problem, we study the effect of satisficing behavior on the transportation systems performance . When network users are satisficing decision-makers, the resulting satisficing user equilibria may degrade the system performance, compared to the perfectly rational user equilibrium. To quantify how much the performance can deteriorate, this paper has quantified the worst-case analytical bound on the price of satisficing. We also quantified the price of satisficing for several networks numerically and compare it to the analytical bound. As we have seen in the numerical examples in this paper, there is a large gap between the worst-case analytical bound and the actual bound. Clearly, this is a limitation of our approach. In the literature of the price of anarchy have similar observations been reported (O'Hare et al., 2016; Colini-Baldeschi et al., 2017). Likewise, the behavior of the price of satisficing in practice can be quite different from what we have observed in this paper. Deriving empirical or network-specific bounds can be meaningful contributions as a future research direction. We suggest additional potential future research directions. For the proposed analytical bound, our result is based on the condition (4.26). By attempting to relax this condition, one may obtain a global bound for any value of κ. In deriving the analytical bound, we utilized a novel technique comparing equilibrium patterns before and after the travel demand is increased, namely $\mathbf{V}$ and $\mathbf{V}_{1+\kappa}$. Applying this technique in the context of the price of risk aversion and the deviation ratio would be an interesting research direction.

References

Ausubel, L.M., Milgrom, P., et al., 2006. The lovely but lonely Vickrey auction. *Combinatorial Auctions* 17, 22–26.

Balas, E., 1985. Disjunctive programming and a hierarchy of relaxations for discrete optimization problems. *SIAM Journal on Algebraic Discrete Methods* 6, 3, 466–486.

Beckmann, M., McGuire, C., Winsten, C., 1956. *Studies in the Economics of Transportation.* Yale University Press, New Haven, Conn.

Bianco, L., Caramia, M., Giordani, S., Piccialli, V., 2015. A game-theoretic approach for regulating hazmat transportation. *Transportation Science* 50, 2, 424–438.

Bonvicini, S., Spadoni, G., 2008. A hazmat multi-commodity routing model satisfying risk criteria: A case study. *Journal of Loss Prevention in the Process Industries* 21, 4, 345–358.

Brown, D.B., Sim, M., 2009. Satisficing Measures for Analysis of Risky Positions. *Management Science* 55, 1, 71–84.

Cantillon, E., Pesendorfer, M., 2006. Auctioning bus routes: The London experience. *Combinatorial auctions*

Charnes, A., Cooper, W.W., 1963. Deterministic equivalents for optimizing and satisficing under chance constraints. *Operations Research* 11, 1, 18–39.

Chen, H.C., Friedman, J.W., Thisse, J.F., 1997. Boundedly rational Nash equilibrium: a probabilistic choice approach. *Games and Economic Behavior* 18, 1, 32–54.

Christodoulou, G., Koutsoupias, E., Spirakis, P.G., 2011. On the performance of approximate equilibria in congestion games. *Algorithmica* 61, 1, 116–140.

Clarke, E.H., 1971. Multipart pricing of public goods. *Public Choice* 11, 1, 17–33.

Colini-Baldeschi, R., Cominetti, R., Mertikopoulos, P., Scarsini, M., 2017. The asymptotic behavior of the price of anarchy. In *International Conference on Web and Internet Economics*, Springer, pp. 133–145.

Dafermos, S., 1980. Traffic equilibrium and variational inequalities. *Transportation Science* 14, 1, 42–54.

Dafermos, S., Nagurney, A., 1984. Sensitivity analysis for the asymmetric network equilibrium problem. *Mathematical Programming* 28, 2, 174–184.

De Vries, S., Vohra, R.V., 2003. Combinatorial auctions: A survey. *INFORMS Journal on Computing* 15, 3, 284–309.

Dempe, S., 2002. *Foundations of Bilevel Programming*. Springer Science & Business Media.

Dempe, S., Mordukhovich, B.S., Zemkoho, A.B., 2014. Necessary optimality conditions in pessimistic bilevel programming. *Optimization* 63, 4, 505–533.

Di, X., He, X., Guo, X., Liu, H.X., 2014. Braess paradox under the boundedly rational user equilibria. *Transportation Research Part B: Methodological* 67, 86–108.

Di, X., Liu, H.X., 2016. Boundedly rational route choice behavior: A review of models and methodologies. *Transportation Research Part B: Methodological* 85, 142–179.

Di, X., Liu, H.X., Pang, J.S., Ban, X.J., 2013. Boundedly rational user equilibria (BRUE): Mathematical formulation and solution sets. *Transportation Research Part B: Methodological* 57, 300–313.

Dolan, E.D., Moré, J.J., 2002. Benchmarking optimization software with performance profiles. *Mathematical Programming* 91, 2, 201–213.

Dunning, I., Huchette, J., Lubin, M., 2017a. JuMP: A modeling language for mathematical optimization. *SIAM Review* 59, 2, 295–320.

Dunning, I., Huchette, J., Lubin, M., 2017b. JuMP: A modeling language for mathematical optimization. *SIAM Review* 59, 2, 295–320.

Eikenbroek, O.A., Still, G.J., van Berkum, E.C., Kern, W., 2018. The boundedly rational user equilibrium: a parametric analysis with application to the network design problem. *Transportation Research Part B: Methodological* 107, 1–17.

Englert, M., Franke, T., Olbrich, L., 2010. Sensitivity of Wardrop equilibria. *Theory of Computing Systems* 47, 1, 3–14.

Esfandeh, T., Kwon, C., Batta, R., 2016. Regulating hazardous materials transportation by dual toll pricing. *Transportation Research Part B: Methodological* 83, 20–35.

Friesz, T.L., Bernstein, D., 2016. *Foundations of network optimization and games*. Springer.

Fujishima, Y., Leyton-Brown, K., Shoham, Y., 1999. Taming the computational complexity of combinatorial auctions: Optimal and approximate approaches. In *IJCAI*, Vol. 99, pp. 548–553.

Gansterer, M., Hartl, R.F., 2017. Collaborative vehicle routing: a survey. *European Journal of Operational Research* 268, 1–12.

Ge, Y., Zhou, X., 2012. An alternative definition of dynamic user optimum on signalised road networks. *Journal of Advanced Transportation* 46, 3, 236–253.

Groves, T., 1973. Incentives in teams. *Econometrica: Journal of the Econometric Society* 41, 4, 617–631.

Guo, X., 2013. Toll sequence operation to realize target flow pattern under bounded rationality. *Transportation Research Part B: Methodological* 56, 203–216.

Guo, X., Liu, H.X., 2011. Bounded rationality and irreversible network change. *Transportation Research Part B: Methodological* 45, 10, 1606–1618.

Han, K., Szeto, W., Friesz, T.L., 2015. Formulation, existence, and computation of boundedly rational dynamic user equilibrium with fixed or endogenous user tolerance. *Transportation Research Part B: Methodological* 79, 16 – 49.

Hara, Y., Hato, E., 2018. A car sharing auction with temporal-spatial OD connection conditions. *Transportation Research Part B: Methodological* 117, 723–739.

Hearn, D.W., Ramana, M.V., 1998. Solving congestion toll pricing models. In *Equilibrium and Advanced Transportation Modelling*. Springer, pp. 109–124.

Hershberger, J., Suri, S., 2001. Vickrey prices and shortest paths: What is an edge worth? In *Foundations of Computer Science, 2001. Proceedings. 42nd IEEE Symposium on*, IEEE, pp. 252–259.

Hicks, R., Madrid, R., Milligan, C., Pruneau, R., Kanaley, M., Dumas, Y., Lacroix, B., Desrosiers, J., Soumis, F., 2005. Bombardier flexjet significantly improves its fractional aircraft ownership operations. *Interfaces* 35, 1, 49–60.

Jaillet, P., Jena, S.D., Ng, T.S., Sim, M., 2016. Satisficing awakens: Models to mitigate uncertainty. *Working Paper*

James, J., Lam, A.Y., Lu, Z., 2018. Double auction-based pricing mechanism for autonomous vehicle public transportation system. *IEEE Transactions on Intelligent Vehicles* 3, 2, 151–162.

Kara, B.Y., Verter, V., 2004. Designing a road network for hazardous materials transportation. *Transportation Science* 38, 2, 188–196.

Kleer, P., Schäfer, G., 2016. The impact of worst-case deviations in non-atomic network routing games. In *International Symposium on Algorithmic Game Theory*, Springer, pp. 129–140.

Koutsoupias, E., Papadimitriou, C., 1999. Worst-case equilibria. In *Annual Symposium on Theoretical Aspects of Computer Science*, Springer, pp. 404–413.

Kuo, A., Miller-Hooks, E., 2015. Combinatorial auctions of railway track capacity in vertically separated freight transport markets. *Journal of Rail Transport Planning & Management* 5, 1, 1–11.

Lam, S.W., Ng, T.S., Sim, M., Song, J.H., 2013. Multiple objectives satisficing under uncertainty. *Operations Research* 61, 1, 214–227.

Lawphongpanich, S., Hearn, D.W., 2004. An MPEC approach to second-best toll pricing. *Mathematical Programming* 101, 1, 33–55.

Lou, Y., Yin, Y., Lawphongpanich, S., 2010. Robust congestion pricing under boundedly rational user equilibrium. *Transportation Research Part B: Methodological* 44, 1, 15–28.

Lozano, L., Smith, J.C., 2017. A value-function-based exact approach for the bilevel mixed-integer programming problem. *Operations Research* 65, 3, 768–786.

Mahmassani, H.S., Chang, G.L., 1987. On boundedly rational user equilibrium in transportation systems. *Transportation Science* 21, 2, 89–99.

Marcotte, P., Mercier, A., Savard, G., Verter, V., 2009. Toll policies for mitigating hazardous materials transport risk. *Transportation Science* 43, 2, 228–243.

Margot, F., 2010. Symmetry in integer linear programming. In *50 Years of Integer Programming 1958-2008.* Springer, pp. 647–686.

Meng, Q., Liu, Z., Wang, S., 2014. Asymmetric stochastic user equilibrium problem with elastic demand and link capacity constraints. *Transportmetrica A: Transport Science* 10, 4, 304–326.

Nakayama, S., Kitamura, R., Fujii, S., 2001. Drivers' route choice rules and network behavior: Do drivers become rational and homogeneous through learning? *Transportation Research Record: Journal of the Transportation Research Board* 1752, 1, 62–68.

Nikolova, E., Stier-Moses, N.E., 2015. The burden of risk aversion in mean-risk selfish routing. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, ACM, pp. 489–506.

Nisan, N., Ronen, A., 2007. Computationally feasible VCG mechanisms. *Journal of Artificial Intelligence Research* 29, 19–47.

O'Hare, S.J., Connors, R.D., Watling, D.P., 2016. Mechanisms that govern how the price of anarchy varies with travel demand. *Transportation Research Part B: Methodological* 84, 55–80.

Östergård, P.R., 2002. A fast algorithm for the maximum clique problem. *Discrete Applied Mathematics* 120, 1-3, 197–207.

de Palma, A., Nesterov, Y., 1998. Optimization formulations and static equilibrium in congested transportation networks. In *CORE Discussion Paper 9861, Université Catholique de Louvain, Louvain-la-Neuve*, pp. 12–17.

Panicucci, B., Pappalardo, M., Passacantando, M., 2007. A path-based double projection method for solving the asymmetric traffic network equilibrium problem. *Optimization Letters* 1, 2, 171–185.

Patriksson, M., 2015. *The traffic assignment problem: models and methods*. Courier Dover Publications.

Pekeč, A., Rothkopf, M.H., 2003. Combinatorial auction design. *Management Science* 49, 11, 1485–1503.

Perakis, G., 2007. The "price of anarchy" under nonlinear and asymmetric costs. *Mathematics of Operations Research* 32, 3, 614–628.

Rassenti, S.J., Smith, V.L., Bulfin, R.L., 1982. A Combinatorial Auction Mechanism for Airport Time Slot Allocation. *The Bell Journal of Economics* 13, 2, 402–417.

Rothkopf, M.H., Pekeč, A., Harstad, R.M., 1998. Computationally manageable combinational auctions. *Management Science* 44, 8, 1131–1147.

Roughgarden, T., Tardos, É., 2002. How bad is selfish routing? *Journal of the ACM (JACM)* 49, 2, 236–259.

Sandholm, T., Likhodedov, A., 2015. Automated design of revenue-maximizing combinatorial auctions. *Operations Research* 63, 5, 1000–1025.

Sandholm, T., Suri, S., Gilpin, A., Levine, D., 2005. CABOB: A fast optimal algorithm for winner determination in combinatorial auctions. *Management Science* 51, 3, 374–390.

Sheffi, Y., 1985. *Urban transportation networks: equilibrium analysis with mathematical programming methods*. title.

Shoup, D.C., 2005. *The High Cost of Free Parking*, Vol. 206. Planners Press Chicago.

Simon, H.A., 1955. A behavioral model of rational choice. *The Quarterly Journal of Economics* pp. 99–118.

Simon, H.A., 1956. Rational choice and the structure of the environment. *Psychological Review* 63, 2, 129.

Smith, M.J., 1979. The existence, uniqueness and stability of traffic equilibria. *Transportation Research Part B: Methodological* 13, 4, 295–304.

Sun, L., Karwan, M.H., Kwon, C., 2017. Generalized bounded rationality and robust multicommodity network design. *Operations Research* 66, 1, 42–57.

Sun, L., Karwan, M.H., Kwon, C., 2018. Generalized bounded rationality and robust multicommodity network design. *Operations Research* 66, 1, 42–57.

Suwansirikul, C., Friesz, T.L., Tobin, R.L., 1987. Equilibrium decomposed optimization: a heuristic for the continuous equilibrium network design problem. *Transportation Science* 21, 4, 254–263.

Szeto, W., Lo, H.K., 2006. Dynamic traffic assignment: properties and extensions. *Transportmetrica* 2, 1, 31–52.

Takalloo, M., Kwon, C., 2018. Sensitivity of Wardrop equilibria: revisited. *Working Paper*

Takalloo, M., Kwon, C., 2019. On the price of satisficing in network user equilibria. *Transportation Science* Accepted.

Tennenholtz, M., 2000. Some tractable combinatorial auctions. In *AAAI/IAAI*, pp. 98–103.

Terry, G.A., 1994. Resort timesharing: Growth and situation analysis. *Journal of Travel & Tourism Marketing* 3, 1, 99–114.

Tsung, C., Ho, H., Lee, S., 2011. An equilibrium-based approach for determining winners in combinatorial auctions. In *Parallel and Distributed Processing with Applications (ISPA), 2011 IEEE 9th International Symposium on*, IEEE, pp. 47–51.

Vickrey, W., 1961. Counterspeculation, auctions, and competitive sealed tenders. *The Journal of Finance* 16, 1, 8–37.

Vorobeychik, Y., Engel, Y., 2011. Average-case analysis of VCG with approximate resource allocation algorithms. *Decision Support Systems* 51, 3, 648–656.

Wächter, A., Biegler, L.T., 2006. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming* 106, 1, 25–57.

Wang, J., Kang, Y., Kwon, C., Batta, R., 2012. Dual toll pricing for hazardous materials transport with linear delay. *Networks and Spatial Economics* 12, 1, 147–165.

Wiesemann, W., Tsoukalas, A., Kleniati, P.M., Rustem, B., 2013. Pessimistic bilevel optimization. *SIAM Journal on Optimization* 23, 1, 353–380.

Wu, J., Sun, H., Wang, D.Z., Zhong, M., Han, L., Gao, Z., 2013. Bounded-rationality based day-to-day evolution model for travel behavior analysis of urban railway network. *Transportation Research Part C: Emerging Technologies* 31, 73–82.

Wu, Q., Hao, J.K., 2015. Solving the winner determination problem via a weighted maximum clique heuristic. *Expert Systems with Applications* 42, 1, 355–365.

Xu, H., Yang, H., Zhou, J., Yin, Y., 2017. A route choice model with context-dependent value of time. *Transportation Science* 51, 2, 536–548.

Ye, H., Yang, H., 2017. Rational behavior adjustment process with boundedly rational user equilibrium. *Transportation Science* 51, 3, 968–980.

Yen, J.Y., 1971. Finding the K shortest loopless paths in a network. *Management Science* 17, 11, 712–716.

Zhang, B., Yao, T., Friesz, T.L., Sun, Y., 2015. A tractable two-stage robust winner determination model for truckload service procurement via combinatorial auctions. *Transportation Research Part B: Methodological* 78, 16–31.

Zhu, S., Levinson, D., 2010. Do people use the shortest path? An empirical test of Wardrop's first principle. In *91th Annual Meeting of the Transportation Research Board, Washington*, Vol. 8.

Appendix A: Copyright Permission for Cahpter 4

USF

Mahdi Takalloo <mtakalloo@mail.usf.edu>

Copyright Permission for Dissertation
9 messages

Mahdi Takalloo <mtakalloo@mail.usf.edu> Mon, Feb 24, 2020 at 4:38 PM
To: permissions@informs.org

to whom it may concern,

This is Mahdi Takalloo, a PhD student at the University of South Florida and an INFORSMS member. One of my papers has been recently accepted for publication in **Transportation Science**. I'd like to ask for copyright permission to use this paper as part of my **PhD dissertation**. The paper title is **On the Price of Satisficing in Network User Equilibria** authored by **Mahdi Takalloo**, and **Changhyun Kwon**.
I appreciate your time and consideration.

Best regards,
Mahdi

--

Mahdi Takalloo,

PhD Student

Department of Industrial and Management Systems Engineering

University of South Florida

Tampa, FL 33620

Matthew Walls <matthew.walls@informs.org> Wed, Feb 26, 2020 at 11:58 AM
To: Mahdi Takalloo <mtakalloo@mail.usf.edu>

Dear Mahdi,

We grant permission at no charge for authors to use their own work in future works, such as dissertations. From your articles online, select **Tools** beneath the **View PDF** button to the left of your article abstract, then **Permissions**, which will open a new web page from our copyright management system. Using that web form, enter the year of publication of the work you want to use, then select **Republish or display content**. Next, select **Republish in a thesis/dissertation**, and complete the requested information about your article, being sure to select **Author of requested content** for the question **Who will republish the content?** You will then be granted permission to republish your article in your dissertation at no charge.

Please let me know if you have any difficulty using the web form.

Best regards,

Matt

informs **Matthew Walls**

Reprint Permission for Chapter 4

Director, Publications

5521 Research Park Drive, Suite 200, Catonsville, MD 21228 USA
p: 443-757-3571 | e: matthew.walls@informs.org

www.informs.org

[Quoted text hidden]

Mahdi Takalloo <mtakalloo@mail.usf.edu> Wed, Feb 26, 2020 at 12:06 PM
To: Matthew Walls <matthew.walls@informs.org>

Dear Matt,

Thanks for your email. My paper is not available in articles online yet. I did the proofread last week. I need to submit my dissertation draft before March 6th. Thanks for your time and consideration.

Best regards,
Mahdi
[Quoted text hidden]

Mahdi Takalloo <mtakalloo@mail.usf.edu> Wed, Feb 26, 2020 at 12:18 PM
To: Matthew Walls <matthew.walls@informs.org>

Dear Matt,

I appreciate it very much if you can guide me through the steps that I need to do to get the permission considering that my article is not available online yet.

Best regards,
Mahdi
[Quoted text hidden]

Matthew Walls <matthew.walls@informs.org> Wed, Feb 26, 2020 at 12:45 PM
To: Mahdi Takalloo <mtakalloo@mail.usf.edu>

Would this email suffice? You retain permission to use your article within your dissertation.

Best,
Matt

[Quoted text hidden]

Mahdi Takalloo <mtakalloo@mail.usf.edu> Wed, Feb 26, 2020 at 12:57 PM
To: Matthew Walls <matthew.walls@informs.org>

Thanks for your email Matt. I will check with the department and get back to you if needed.

Best regards,
Mahdi
[Quoted text hidden]

Reprint Permission for Chapter 4

Mahdi Takalloo <mtakalloo@mail.usf.edu> Tue, Mar 24, 2020 at 3:27 PM
To: "Schmidt, Leetta" <lmschmidt@usf.edu>

Hi Leetta,

Please find my emails with INFORMS publication director.

Best regards,
Mahdi
[Quoted text hidden]

Schmidt, Leetta <lmschmidt@usf.edu> Tue, Mar 24, 2020 at 3:38 PM
To: Mahdi Takalloo <mtakalloo@mail.usf.edu>

Hello Mahdi,

Thank you for sending the journal information and forwarding this communication. Yes, the email below from the INFORMS publication director will suffice as the permissions required by the journal's copyright transfer agreement (https://pubsonline.informs.org/pb-assets/copyright-TRSC-8-2017-form.pdf). You can save a copy/printout of this email chain in your dissertation's appendix (at least enough to identify your article and capture his permissions).

Please let me know if you have any further questions.

Best,

LeEtta

LeEtta Schmidt

Copyright and Intellectual Property Librarian

USF Libraries

University of South Florida

813-974-1627

[Quoted text hidden]

Mahdi Takalloo <mtakalloo@mail.usf.edu> Tue, Mar 24, 2020 at 3:45 PM
To: "Schmidt, Leetta" <lmschmidt@usf.edu>

Thank you very much.
[Quoted text hidden]

Reprint Permission for Chapter 4

Appendix B: Proofs of Chapter 2

Proof of Proposition 2.1. Formulations (P1) and (P2) possess the same objective function. Moreover, constraints (2.12) and (2.17) are exactly the same. Therefore, we only need to show that (2.18) is equivalent to (2.13)–(2.14).

1. We show that if either (2.13) or (2.14) is violated, then (2.18) is also violated. First, suppose that $\overline{\mathbf{x}}$ violates (2.13); i.e., there exist j, v, n, and m such that

$$e_m \overline{x}_{jv} > s_n + M(1 - \overline{x}_{lv}) \tag{A1}$$

$$s_m \leq s_n \leq e_m, \tag{A2}$$

where we abbreviated the index descriptions, $i, k \in \mathcal{I}, j \in \mathcal{B}_i, l \in \mathcal{B}_k, m \in \mathcal{T}_j, n \in \mathcal{T}_l$, for simplicity. For (A1), $\overline{x}_{jv} = \overline{x}_{lv} = 1$ is the only possibility. From (A2), we observe:

$$s_n \leq e_m \implies s_n \leq e_m + r_{ike_m s_n}$$

$$s_m \leq s_n \implies s_m \leq s_n \leq e_n \leq e_n + r_{ike_n s_m}$$

Therefore, when (A2) holds, (2.20) also holds. In this case, $\overline{\mathbf{x}}$ does not satisfy (2.18).

Similarly, when $\overline{\mathbf{x}}$ violates (2.14), we can show that $\overline{\mathbf{x}}$ also violates (2.18).

2. We show that if (2.18) is violated, then (2.13)–(2.14) are also violated. Suppose that $\overline{\mathbf{x}}$ violates (2.18); i.e., there exist j, v, n, and m such that

$$\overline{x}_{jv} = \overline{x}_{lv} = 1,$$

$$s_m \leq e_n + r_{ike_n s_m},$$

$$s_n \leq e_m + r_{ike_m s_n}.$$

If $s_n \leq s_m$, we obtain $s_n \leq s_m \leq e_m \leq e_m + r_{ike_m s_n}$, and as a result $\overline{\mathbf{x}}$ violates (2.13)–(2.14). If $s_m \leq s_n$, then $s_m \leq s_n \leq e_n \leq e_n + r_{ike_n s_m}$, and consequently, $\overline{\mathbf{x}}$ violates (2.13)–(2.14).

Hence, (2.18) is equivalent to (2.13)–(2.14). □

Proof of Proposition 2.2. Since we have p bids, there are $\binom{p}{2}$ pairs of possible bids. Moreover, since the maximum number of trips in a submitted bid is q, vector L has a maximum possible length of $2q$. Therefore, the time complexity of sorting L is $O(q \log q)$. After sorting $O(q)$ comparisons should be made. Consequently, the time complexity of Algorithm 2.2 is

$$\binom{p}{2}[O(2q \log q) + O(q)].$$

Hence, the result follows. □

Proof of Proposition 2.3. Let $Z^{*,v}_{\mathcal{V}}$ be the social welfare associated with the v-th vehicle; that is, $Z^{*,v}_{\mathcal{V}} = \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{B}_i} c_j x^*_{jv}$ where $\mathbf{x}^*$ is an optimal solution to (P2) with $\mathcal{V}$. Without loss of generality, let us assume that $Z^{*,v}_{\mathcal{V}} \geq Z^{*,v'}_{\mathcal{V}}$ if $v \leq v'$. Since $Z^*_{\mathcal{V}'}$ is optimal, we have

$$Z^*_{\mathcal{V}'} \geq \sum_{v=1}^{|\mathcal{V}'|} Z^{*,v}_{\mathcal{V}}$$

or equivalently

$$\frac{Z^*_{\mathcal{V}'}}{|\mathcal{V}'|} \geq \frac{\sum_{v=1}^{|\mathcal{V}'|} Z^{*,v}_{\mathcal{V}}}{|\mathcal{V}'|} \tag{A3}$$

Moreover, since $Z^{*,v}_{\mathcal{V}} \geq Z^{*,v'}_{\mathcal{V}}$ if $v \leq v'$, we have

$$\frac{Z^*_{\mathcal{V}'}}{|\mathcal{V}'|} \geq \frac{\sum_{v=1}^{|\mathcal{V}'|} Z^{*,v}_{\mathcal{V}}}{|\mathcal{V}'|} \geq \frac{\sum_{v=|\mathcal{V}'|+1}^{|\mathcal{V}|} Z^{*,v}_{\mathcal{V}}}{|\mathcal{V}| - |\mathcal{V}'|} \tag{A4}$$

Using the weighted average of the right-hand-side of (A3) and (A4), we obtain the following:

$$\frac{Z^*_{\mathcal{V}'}}{|\mathcal{V}'|} \geq \frac{|\mathcal{V}'|}{|\mathcal{V}|} \frac{\sum_{v=1}^{|\mathcal{V}'|} Z^{*,v}_{\mathcal{V}}}{|\mathcal{V}'|} + \frac{|\mathcal{V}| - |\mathcal{V}'|}{|\mathcal{V}|} \frac{\sum_{v=1+|\mathcal{V}'|}^{|\mathcal{V}|} Z^{*,v}_{\mathcal{V}}}{|\mathcal{V}| - |\mathcal{V}'|}$$

$$= \frac{\sum_{v=1}^{|\mathcal{V}|} Z_{\mathcal{V}}^{*,v}}{|\mathcal{V}|} = \frac{Z_{\mathcal{V}}^{*}}{|\mathcal{V}|} \tag{A5}$$

Furthermore, from the sub-optimality of SSVD, we have

$$Z_{\mathcal{V}}^{G} \leq Z_{\mathcal{V}}^{*},$$

which along with (A5) leads to

$$Z_{\mathcal{V}}^{*} - Z_{\mathcal{V}}^{G} \leq \frac{|\mathcal{V}|}{|\mathcal{V}'|} Z_{\mathcal{V}'}^{*} - Z_{\mathcal{V}}^{G}.$$

This completes the proof. □

Proof of Proposition 2.4. Suppose $\overline{\mathbf{x}}$ is a feasible point for problem (P3). We let $\overline{y}_j = \sum_{v\in\mathcal{V}} \overline{x}_{jv}$. From (2.23) we have

$$\sum_{j\in\mathcal{B}_i} \sum_{v\in\mathcal{V}} \overline{x}_{jv} = \sum_{j\in\mathcal{B}_i} \overline{y}_j \leq 1.$$

Getting summation over v in (2.24) we have

$$\sum_{v\in\mathcal{V}} \sum_{j\in\mathcal{C}_m} \overline{x}_{jv} = \sum_{j\in\mathcal{C}_m} \overline{y}_j \leq |\mathcal{V}|.$$

Therefore, $\overline{\mathbf{x}}$ is a feasible solution for problem (R). Moreover, (P3) and (R) share the same objective function:

$$\sum_{j\in\mathcal{B}} \sum_{v\in\mathcal{V}} c_j x_{jv} = \sum_{j\in\mathcal{B}} c_j y_j$$

Hence, problem (R) is a relaxation of problem (P3). □

Proof of Proposition 2.5. Under incentive compatibility, we have

$$v_i = Z_{\text{WDP}}^{\epsilon} - \sum_{j\in\mathcal{B}_{-i}} c_j x_j^{\epsilon}, \quad \forall i \in \mathcal{I}.$$

Since $Z^{\epsilon}_{\text{WDP}} \geq Z^{\epsilon}_{\text{WDP}_{-i}}$, we obtain the following:

$$v_i = Z^{\epsilon}_{\text{WDP}} - \sum_{j \in \mathcal{B}_{-i}} c_j x^{\epsilon}_j \geq Z^{\epsilon}_{\text{WDP}_{-i}} - \sum_{j \in \mathcal{B}_{-i}} c_j x^{\epsilon}_j \geq p^{\epsilon}_i, \quad \forall i \in \mathcal{I}.$$

Since $v_i \geq p^{e}_i$, the bidders do not lose by participating in the auction, and therefore rationality holds. □

Proof of Proposition 2.6. Considering VCG payment in (2.5), we have

$$p^*_i = Z^*_{\text{WDP}_{-i}} - \left(Z^0_{\text{WDP}} - \sum_{j \in \mathcal{B}_i} c_j x^0_j \right) \tag{A6}$$

$$p^{\epsilon}_i = Z^{\epsilon}_{\text{WDP}_{-i}} - \left(Z^{\epsilon}_{\text{WDP}} - \sum_{j \in \mathcal{B}_i} c_j x^{\epsilon}_j \right). \tag{A7}$$

From (A6) and (A7) we obtain

$$p^*_i - p^{\epsilon}_i = (Z^*_{\text{WDP}_{-i}} - Z^{\epsilon}_{\text{WDP}_{-i}}) + (Z^{\epsilon}_{\text{WDP}} - Z^*_{\text{WDP}}) + \left(\sum_{j \in \mathcal{B}_i} c_j x^*_j - \sum_{j \in \mathcal{B}_i} c_j x^{\epsilon}_j \right). \tag{A8}$$

Since the WDP is a maximization problem, we have the following:

$$Z^{\epsilon}_{\text{WDP}} \leq Z^*_{\text{WDP}}. \tag{A9}$$

Since at most one bid from each bidder can be determined as a winner, we have

$$\left(\sum_{j \in \mathcal{B}_i} c_j x^*_j - \sum_{j \in \mathcal{B}_i} c_j x^{\epsilon}_j \right) \leq \max_{j \in \mathcal{B}_i} c_j. \tag{A10}$$

Moreover, from (2.30), we have

$$\frac{Z^*_{\text{WDP}} - Z^{\epsilon}_{\text{WDP}}}{Z^{\epsilon}_{\text{WDP}}} \leq \epsilon \tag{A11}$$

Substituting (A9)–(A11) to (A8), we obtain:

$$p_i^* - p_i^\epsilon \leq \epsilon Z^\epsilon_{\text{WDP}_{-i}} + \max_{j \in \mathcal{B}_i} c_j,$$
$$p_i^\epsilon - p_i^* \leq \epsilon Z^\epsilon_{\text{WDP}} + \max_{j \in \mathcal{B}_i} c_j.$$

Consequently,

$$|p_i^\epsilon - p_i^*| \leq \epsilon \max\left\{Z^\epsilon_{\text{WDP}}, Z^\epsilon_{\text{WDP}_{-i}}\right\} + \max_{j \in \mathcal{B}_i} c_j,$$

which completes the first part of the proof.

The revenue of the auctioneer is the summation of the bidder's payments. Hence, we have

$$R^* = \sum_{i \in \mathcal{I}} p_i^* = \sum_{i \in \mathcal{I}} Z^*_{\text{WDP}_{-i}} - |\mathcal{I}| Z^*_{\text{WDP}} + \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{B}_i} c_j x_j^* = \sum_{i \in \mathcal{I}} Z^*_{\text{WDP}_{-i}} - (|\mathcal{I}| - 1) Z^*_{\text{WDP}}.$$

Similarly, we can write R^ϵ as

$$R^\epsilon = \sum_{i \in \mathcal{I}} Z^\epsilon_{\text{WDP}_{-i}} - (|\mathcal{I}| - 1) Z^\epsilon_{\text{WDP}}.$$

Using (A9) and (A11), we obtain

$$R^\epsilon - R^* \leq (|\mathcal{I}| - 1)\epsilon Z^\epsilon_{\text{WDP}}$$
$$R^* - R^\epsilon \leq \epsilon \sum_{i \in \mathcal{I}} Z^\epsilon_{\text{WDP}_{-i}}.$$

Consequently,

$$|R^\epsilon - R^*| \leq \epsilon \max\left\{(|\mathcal{I}| - 1) Z^\epsilon_{\text{WDP}}, \sum_{i \in \mathcal{I}} Z^\epsilon_{\text{WDP}_{-i}}\right\},$$

which completes the proof. □

Appendix C: Proofs of Chapter 3

Proof Proposition 3.1. The set of feasible solution is non-empty and bounded, since there always exists a feasible solution, for example, with $\mathbf{t} = \mathbf{0}$ and the toll on each link can be bounded (Marcotte et al., 2009). Therefore, the existence of an optimal solution is guaranteed by Dempe (2002, Theorem 3.3). The non-existence of optimal solution to PHTP is shown by the example in Section 3.4. $\square$

Proof Lemma 3.1. We denote the feasible region of $\overline{\text{PHTP}_\epsilon}$ by:

$$\mathbf{T}_\epsilon = \left\{ (\mathbf{t}, z) : \sum_{s \in \mathcal{S}} f_s(\mathbf{t}, \mathbf{x}_s) - z \leq 0, \quad \forall \mathbf{x} \in \mathbb{X}^\epsilon(\mathbf{t}), \quad \mathbf{0} \leq \mathbf{t} \leq \mathbf{B} \right\}.$$

We show that the feasible set $\mathbf{T}_\epsilon$ is closed for all $\epsilon \geq \mathbf{0}$, following the proof of Wiesemann et al. (2013, Proposition 3.1). Although our lower-level problem $\mathbb{X}^\epsilon(\cdot)$ is discrete unlike the problem considered in Wiesemann et al. (2013), we have an equivalent continuous LP relaxation. With this slight modification, $\mathbf{T}_\epsilon$ can be shown closed. In addition, $\mathbf{T}_\epsilon$ is obviously non-empty; e.g. $\mathbf{t} = \mathbf{0}$ generates a feasible solution. Since $\overline{\text{PHTP}_\epsilon}$ minimizes a continuous objective function over a closed, non-empty set, there exists an optimal solution by the extreme value theorem. $\square$

Proof of Proposition 3.2. Consider toll vector $\widehat{\mathbf{t}} \geq \mathbf{0}$ where $\widehat{t}_{ij} > B_{ij}$ for some $(i, j) \in \mathcal{A}$. Let $\widehat{\mathbf{x}}$ be a solution to the lower level problem under $\widehat{\mathbf{t}}$. Note that $(\widehat{\mathbf{t}}, \widehat{\mathbf{x}})$ is feasible to PHTP_ϵ but not $\overline{\text{PHTP}_\epsilon}$. We consider two cases.

(i) Consider cases when $h_s(\widehat{\mathbf{t}}, \widehat{\mathbf{x}}_s) > (B_{ij} - \max_{s \in \mathcal{S}} \epsilon_s)$ for some $s \in \mathcal{S}$. That is,

$$h_s(\widehat{\mathbf{t}}, \widehat{\mathbf{x}}_s) = \sum_{(i,j) \in \mathcal{A}} (c_{ij} + \widehat{t}_{ij} + \beta \rho_{ij}) \widehat{x}_{sij}$$

$$
\begin{aligned}
&> B_{ij} - \max_{s \in \mathcal{S}} \epsilon_s \\
&= \frac{z_{\mathbf{t=0}}}{\alpha} + \beta \sum_{(i,j) \in \mathcal{A}} \rho_{ij},
\end{aligned}
$$

which leads to

$$
\sum_{(i,j) \in \mathcal{A}} (c_{ij} + \widehat{t}_{ij}) \widehat{x}_{sij} > \frac{z_{\mathbf{t=0}}}{\alpha}.
$$

Then we have:

$$
\begin{aligned}
f_s(\widehat{\mathbf{t}}, \widehat{\mathbf{x}}) &= \sum_{(i,j) \in \mathcal{A}} n_s \Big(\rho_{ij} + \alpha (c_{ij} + \widehat{t}_{ij}) \Big) \widehat{x}_{sij} \\
&\geq \sum_{(i,j) \in \mathcal{A}} n_s \Big(\alpha (c_{ij} + \widehat{t}_{ij}) \Big) \widehat{x}_{sij} \\
&\geq \sum_{(i,j) \in \mathcal{A}} \alpha \Big(c_{ij} + \widehat{t}_{ij} \Big) \widehat{x}_{sij} \\
&> z_{\mathbf{t=0}}.
\end{aligned}
$$

Consequently $z_{\widehat{\mathbf{t}}} > z_{\mathbf{t=0}}$. Therefore, $\mathbf{t} = \mathbf{0}$ is a better solution that is feasible to $\overline{\text{PHTP}_\epsilon}$.

(ii) Consider cases when $h_s(\widehat{\mathbf{t}}, \widehat{\mathbf{x}}) \leq (B_{ij} - \max_{s \in \mathcal{S}} \epsilon_s)$ for all $s \in \mathcal{S}$. That is,

$$
\sum_{(i,j) \in \mathcal{A}} (c_{ij} + \widehat{t}_{ij} + \beta \rho_{ij}) \widehat{x}_{sij} = h_s(\widehat{\mathbf{t}}, \widehat{\mathbf{x}}_s) \leq B_{ij} - \max_{s \in \mathcal{S}} \epsilon_s.
$$

Note $\widehat{t}_{ij} \widehat{x}_{sij} \leq h_s(\widehat{\mathbf{t}}, \widehat{\mathbf{x}}_s)$ for all $(i,j) \in \mathcal{A}$. Therefore, we have $\widehat{t}_{ij} \leq (B_{ij} - \max_{s \in \mathcal{S}} \epsilon_s)$ for all $(i,j) \in \mathcal{A}$ such that $\widehat{x}_{sij} = 1$. Construct $\bar{\mathbf{t}}$, which is feasible to $\overline{\text{PHTP}_\epsilon}$, as follows:

$$
\bar{t}_{ij} = \begin{cases} \widehat{t}_{ij} & \text{if } \widehat{t}_{ij} \leq B_{ij}, \\ B_{ij} & \text{if } \widehat{t}_{ij} > B_{ij}, \end{cases}
$$

for all $(i, j) \in \mathcal{A}$. Since $h_s(\widehat{\mathbf{t}}, \widehat{\mathbf{x}}_s) \leq (B_{ij} - \max_{s \in \mathcal{S}} \epsilon_s)$, we also have $h_s(\widehat{\mathbf{t}}, \widehat{\mathbf{x}}_s) + \epsilon_s \leq B_{ij}$. Therefore, the following must hold:

$$\widehat{x}_{sij} = \begin{cases} 0 \text{ or } 1 & \text{if } \widehat{t}_{ij} < B_{ij}, \\ 0 & \text{if } \widehat{t}_{ij} \geq B_{ij}, \end{cases}$$

Moreover, based on the definition of $\bar{\mathbf{t}}$, the following also holds:

$$\widehat{x}_{sij} = \begin{cases} 0 \text{ or } 1 & \text{if } \bar{t}_{ij} < B_{ij}, \\ 0 & \text{if } \bar{t}_{ij} = B_{ij}, \end{cases}$$

Since $\widehat{t}_{ij} = \bar{t}_{ij}$ for all $(i, j) \in \mathcal{A}$ where $\widehat{t}_{ij} \leq B_{ij}$, we have:

$$\begin{aligned} h_s(\widehat{\mathbf{t}}, \widehat{\mathbf{x}}_s) &= \sum_{(i,j) \in \mathcal{A}} (c_{ij} + \widehat{t}_{ij} + \beta \rho_{ij}) \widehat{x}_{sij} \\ &= \sum_{(i,j) \in \mathcal{A}: \widehat{t}_{ij} < B_{ij}} (c_{ij} + \widehat{t}_{ij} + \beta \rho_{ij}) \widehat{x}_{sij} + \sum_{(i,j) \in \mathcal{A}: \widehat{t}_{ij} \geq B_{ij}} (c_{ij} + \widehat{t}_{ij} + \beta \rho_{ij}) \widehat{x}_{sij} \\ &= \sum_{(i,j) \in \mathcal{A}: \bar{t}_{ij} < B_{ij}} (c_{ij} + \bar{t}_{ij} + \beta \rho_{ij}) \widehat{x}_{sij} + \sum_{(i,j) \in \mathcal{A}: \bar{t}_{ij} = B_{ij}} (c_{ij} + \bar{t}_{ij} + \beta \rho_{ij}) \widehat{x}_{sij} \\ &= \sum_{(i,j) \in \mathcal{A}} (c_{ij} + \bar{t}_{ij} + \beta \rho_{ij}) \widehat{x}_{sij} \\ &= h_s(\bar{\mathbf{t}}, \widehat{\mathbf{x}}_s) \end{aligned}$$

Therefore, the new pair $(\bar{\mathbf{t}}, \widehat{\mathbf{x}})$ is still feasible to the problem. Note that, however, $z_{\widehat{\mathbf{t}}} \geq z_{\bar{\mathbf{t}}}$, since the toll values have been decreased.

In summary, for any solution $\widehat{\mathbf{t}}$ that is feasible to PHTP_ϵ, but not $\overline{\text{PHTP}_\epsilon}$, there exists a solution $\bar{\mathbf{t}}$ to $\overline{\text{PHTP}_\epsilon}$ such that $z_{\widehat{\mathbf{t}}} \geq z_{\bar{\mathbf{t}}}$. Therefore, we obtain the proposition. □

Proof of Proposition 3.3. Consider a feasible solution $(\mathbf{t}^k, \mathbf{z}^k, \mathbf{y}^k)$ where $\mathbf{y}^k$ involves a subtour for some $s \in \mathcal{S}$. After eliminating subtours from $\mathbf{y}^k$, we obtain $\overline{\mathbf{y}}$. We suppose that $\mathbf{t}^k$ and $\overline{\mathbf{y}}$ induce $\overline{\mathbf{z}}$. We observe that

$$h_s(\mathbf{t}^k, \mathbf{x}_s^l) \geq h_s(\mathbf{t}^k, \mathbf{y}_s^k) + \epsilon_s > h_s(\mathbf{t}^k, \overline{\mathbf{y}}_s) + \epsilon_s,$$

which implies that $\overline{\mathbf{y}}$ will create less constraints of $[f_s(\mathbf{t}, \mathbf{x}_s^l) - z_s \leq 0]$ types. Therefore $\sum_{s\in\mathcal{S}} \widehat{z}_s \geq \sum_{s\in\mathcal{S}} \overline{z}_s$; that is, solutions without subtours always generate a better or the same objective function value. □

Proof of Lemma 3.2. Note that the optimal objective function value $\sum_{s\in\mathcal{S}} z_s^k$ can be divided into two components: the toll independent part $\sum_{s\in\mathcal{S}} \sum_{(i,j)\in\mathcal{A}} (n_s\rho_{ij} + \alpha n_s c_{ij}) x_{sij}^{l(s)}$ and the toll dependent component $\alpha \sum_{s\in\mathcal{S}} \sum_{(i,j)\in\mathcal{A}} n_s t_{ij}^k x_{sij}^{l(s)}$. Since $R^{\mathbf{0}}$ is the minimum possible value for the toll independent component, the following inequality holds:

$$R^{\mathbf{0}} \leq \sum_{s\in\mathcal{S}} \sum_{(i,j)\in\mathcal{A}} (n_s\rho_{ij} + \alpha n_s c_{ij}) x_{sij}^{l(s)}.$$

Consequently, we have

$$\begin{aligned} R^{\mathbf{0}} + \alpha \sum_{s\in\mathcal{S}} n_s \sum_{(i,j)\in\mathcal{A}} t_{ij}^k x_{sij}^l &= R^{\mathbf{0}} + \alpha \sum_{s\in\mathcal{S}} n_s T_s^k \\ &\leq \sum_{s\in\mathcal{S}} \sum_{(i,j)\in\mathcal{A}} n_s\big(\rho_{ij} + \alpha(c_{ij} + t_{ij}^k)\big) x_{sij}^{l(s)} = \sum_{s\in\mathcal{S}} z_s^k. \end{aligned} \tag{A12}$$

From (3.11), we have:

$$\sum_{s\in\mathcal{S}} z_s^k \leq U_{\overline{\mathbf{t}}}. \tag{A13}$$

From (A12) and (A13), we obtain the lemma. □

Proof of Lemma 3.3. From (3.13) we obtain:

$$h_s(\mathbf{t}^k, \mathbf{y}_s^k) \geq h_s(\mathbf{t}^k, \widehat{\mathbf{y}}(\mathbf{t}^k)_s) \qquad \forall s \in \mathcal{S}$$

Therefore (3.8b) holds. Moreover, as $\widehat{\mathbf{y}}(\mathbf{t}^k) \in \mathbf{X}$, we know that $(\mathbf{t}^k, \mathbf{z}^k, \widehat{\mathbf{y}}^k)$ is feasible to Master$_k$. Since $(\mathbf{t}^k, \mathbf{z}^k, \widehat{\mathbf{y}}(\mathbf{t}^k))$ and $(\mathbf{t}^k, \mathbf{z}^k, \mathbf{y}^k)$ have the same objective function value of $\sum_{s\in\mathcal{S}} z_s^k$, we observe that $(\mathbf{t}^k, \mathbf{z}^k, \widehat{\mathbf{y}}(\mathbf{t}^k))$ is also optimal to Master$_k$. □

Proof of Proposition 3.4. From (3.14) we obtain:

$$\sum_{(i,j)\in\mathcal{A}} (c_{ij} + \beta\rho_{ij}) x_{sij}^{l(s)} \leq C_s^0 \qquad \forall s \in \mathcal{S} \tag{A14}$$

Additionally, based on Lemma 3.2, we have:

$$\sum_{(i,j)\in\mathcal{A}} t_{ij}^k x_{sij}^{l(s)} \leq \overline{T}_s \qquad \forall s \in \mathcal{S} \tag{A15}$$

Consequently, for any ϵ-approximation solution in $\mathbf{X}^k$ for each $s \in \mathcal{S}$, we have:

$$h_s(\mathbf{t}^k, \mathbf{x}_s^l) < h_s(\mathbf{t}^k, \mathbf{x}_s^{l(s)}) + \epsilon_s \leq C_s^0 + \overline{T}_s + \epsilon_s \leq \max_{s'\in\mathcal{S}} \{C_{s'}^0 + \overline{T}_{s'} + \epsilon_{s'}\} \quad \forall s \in \mathcal{S}, l \in [\![|\mathbf{X}^k|]\!].$$

Let us define $\mathcal{A}(\mathbf{t}^k) = \{(i,j) \mid t_{ij}^k > \max_{s'\in\mathcal{S}} \{C_{s'}^0 + \overline{T}_{s'} + \epsilon_{s'}\} - \beta\rho_{ij} - c_{ij}\}$. Consider the toll vector $\widetilde{\mathbf{t}}^k$ defined as follows:

$$\widetilde{t}_{ij}^k = \min\left\{t_{ij}^k,\ \max_{s'\in\mathcal{S}} \{C_{s'}^0 + \overline{T}_{s'} + \epsilon_{s'}\} - \beta\rho_{ij} - c_{ij}\right\}$$

for each $(i,j) \in \mathcal{A}$. That is,

$$\widetilde{t}_{ij}^k = \begin{cases} t_{ij}^k & \text{if } (i,j) \notin \mathcal{A}(\mathbf{t}^k), \\ \max_{s'\in\mathcal{S}} \{C_{s'}^0 + \overline{T}_{s'} + \epsilon_{s'}\} - \beta\rho_{ij} - c_{ij} & \text{if } (i,j) \in \mathcal{A}(\mathbf{t}^k). \end{cases}$$

Suppose the path represented by $\mathbf{x}_s^l$ passes through arcs in $\mathcal{A}(\mathbf{t}^k)$; i.e. $x_{sij}^l = 1$ for some $(i,j) \in \mathcal{A}(\mathbf{t}^k)$. Then, we have:

$$h_s(\mathbf{t}^k, \mathbf{x}_s^l) \geq \max_{s' \in \mathcal{S}} \{C_{s'}^0 + \overline{T}_{s'} + \epsilon_{s'}\},$$
$$h_s(\widetilde{\mathbf{t}}^k, \mathbf{x}_s^l) \geq \max_{s' \in \mathcal{S}} \{C_{s'}^0 + \overline{T}_{s'} + \epsilon_{s'}\}.$$

Thus, the path represented by $\mathbf{x}_s^l$ is not an ϵ_s-*approximate* path under both $\mathbf{t}^k$ and $\widetilde{\mathbf{t}}^k$. Now suppose the path represented by $\mathbf{x}_s^l$ does not pass through arc $(i,j) \in \mathcal{A}(\mathbf{t}^k)$; i.e. $x_{sij}^l = 0$ for all $(i,j) \in \mathcal{A}(\mathbf{t}^k)$. Then we have:

$$h_s(\mathbf{t}^k, \mathbf{x}_s^l) = h_s(\widetilde{\mathbf{t}}^k, \mathbf{x}_s^l). \tag{A16}$$

Consequently, based on (A16), all ϵ-*approximate* paths under toll $\mathbf{t}^k$ are ϵ-*approximate* paths under $\widetilde{\mathbf{t}}^k$. Moreover, all non-ϵ-*approximate* paths under toll $\mathbf{t}^k$ are non-ϵ-*approximate* paths under $\widetilde{\mathbf{t}}^k$.

Since ϵ-*approximate* paths do not include any arc $(i,j) \in \mathcal{A}(\mathbf{t}^k)$, the toll amount charged on the satisficing paths do not change by setting the toll to $\widetilde{\mathbf{t}}^k$. Hence, if $(\mathbf{t}^k, \mathbf{z}^k, \mathbf{y}^k)$ is optimal to Master$_k$, so is $(\widetilde{\mathbf{t}}^k, \mathbf{z}^k, \mathbf{y}^k)$. Therefore, M_{sij}^3 can be bounded by (3.15). □

Proof of Lemma 3.5. Suppose not. In particular, suppose there exists a shipment $s \in \mathcal{S}$ for which there is no ϵ-*approximate* path. Consequently, based on (3.8b), we have $z_s^k = 0$. We pick $\mathbf{x}_s = \arg\min_{\mathbf{y}_s \in \mathbf{X}_s} \ h_s(\mathbf{t}^k, \mathbf{y}_s)$ in Sub$_k$. Then we have

$$w_s \leq h_s^k - h_s(\mathbf{t}^k, \mathbf{x}_s) + \epsilon_s = \epsilon_s$$
$$w_s \leq f_s(\mathbf{t}^k, \mathbf{x}_s) - z_s^k = f_s(\mathbf{t}^k, \mathbf{x}_s).$$

Since Sub$_k$ is a maximization problem, we have: $w_s^k = \min\{\epsilon_s, f_s(\mathbf{t}^k, \mathbf{x}_s^k)\} > 0$, which is a contradiction. □

Proof of Proposition 3.5. Based on Lemma 3.3, $(\mathbf{t}^k, \mathbf{z}^k, \widehat{\mathbf{y}}(\mathbf{t}^k), \mathbf{v}^k, \widehat{\boldsymbol{\tau}})$ is also optimal to Master$_k$, where $\widehat{\mathbf{y}}(\mathbf{t}^k)$ is optimal to (3.13). We have:

$$\sum_{(i,j)\in\mathcal{A}} (c_{ij} + \beta\rho_{ij} + t^k_{ij}) x^{l(s)}_{sij} + \epsilon_s \geq \sum_{(i,j)\in\mathcal{A}} (c_{ij} + \beta\rho_{ij} + t^k_{ij}) \widehat{y}_{sij} + \epsilon_s \qquad \forall s \in \mathcal{S}.$$

From (A14) and (A15), we have:

$$\sum_{(i,j)\in\mathcal{A}} (c_{ij} + t^k_{ij} + \beta\rho_{ij}) x^{l(s)}_{sij} \leq C^0_s + \overline{T}_s + \epsilon_s \quad \forall s \in \mathcal{S}.$$

Consequently:

$$\sum_{(i,j)\in\mathcal{A}} (c_{ij} + t^k_{ij} + \beta\rho_{ij}) \widehat{y}_{sij}(\mathbf{t}^k) + \epsilon_s \leq C^0_s + \overline{T}_s + \epsilon_s \qquad \forall s \in \mathcal{S}$$

which completes the proof for M^{1l}_s. By substituting t_{ij} with (3.15) in (3.16c), we can easily bound M^{2l}_s. □

Proof of Lemma 3.5. Given a disjunctive set $\bigcup_{i=1}^d P^i$, where each $P^i = \{x \in \mathbb{R}^n : \mathbf{a}^i\mathbf{x} \leq b_i\}$ is a bounded polyhedra, the formulation of Balas (1985) is

$$\mathbf{x} = \sum_{i=1}^{d} \mathbf{x}^i \tag{A17a}$$

$$\mathbf{a}^i\mathbf{x}^i \leq b_i u_i \qquad \forall i \in [\![d]\!] \tag{A17b}$$

$$\sum_{i=1}^{d} u_i = 1 \tag{A17c}$$

$$u \in \{0,1\}^d. \tag{A17d}$$

where the $\mathbf{x}^i$ are auxiliary continuous variables that serve as copies of x for each disjunct, and u are binary variables that indicate each of the possible disjuncts P^i. In our setting, $d = 2$, and we take the copies of the original decision variables $(\mathbf{t}_s, z_s, \boldsymbol{\tau}_s, \mathbf{y}_s)$ corresponding to the disjuncts P_s^{1l} and P_2^{2l} to be $(\mathbf{t}_s^{1l}, z_s^{1l}, \boldsymbol{\tau}_s^{1l}, \mathbf{y}_s^{1l})$ and $(\mathbf{t}_s^{1l}, z_s^{1l}, \boldsymbol{\tau}_s^{1l}, \mathbf{y}_s^{1l})$, respectively.

We can observe that (3.20a) and (3.20b) derive from (A17b). Similarly, (3.20c) and (3.20d) also derive from (A17b) by homogenizing all constant terms that appear in the constraints describing $Q_s(1)$. Finally, (3.20e), (3.20f), and (3.20g) are analogous to (A17a), (A17c), and (A17d), respectively □

Proof of Proposition 3.6. Follows directly from applying Lemma 3.5 for each $s \in \mathcal{S}$ and $l \in [\![|\mathbf{X}^k|]\!]$. Note that we have explicitly written out the constraints describing the sets $Q_s(u_s^{1l})$. Finally, we explicitly impose the integrality on the y variables relaxed in the description of $Q_s(\cdot)$ in constraints (3.21q). □

Appendix D: Proofs of Chapter 4

Proof of Lemma 4.1. Given $\overline{\mathbf{f}}$, we let $\overline{\mathbf{v}}$ be the arc flow vector from $\overline{\mathbf{f}} \mapsto \overline{\mathbf{v}}$. Let $\overline{\epsilon}$ be the perception error that makes $\overline{\mathbf{x}}$ a solution to (4.12). Then, $\overline{\mathbf{x}}$ is a user equilibrium flow with respect to arc travel time $\lambda_a^w t_a(\cdot)$ and the following follows from (4.4):

$$\overline{f}_p > 0 \implies \sum_{a \in \mathcal{A}} \delta_a^p \overline{\lambda}_a^w t_a(\overline{\mathbf{v}}) = \min_{p' \in \mathcal{P}_w} \sum_{a \in \mathcal{A}} \delta_a^{p'} \overline{\lambda}_a^w t_a(\overline{\mathbf{v}}) \tag{A18}$$

for all $p \in \mathcal{P}_w$ and $w \in \mathcal{W}$. Since $\overline{\lambda}_a^w \in [\frac{1}{1+\kappa}, 1]$, the right-hand-side of (A18) implies

$$\frac{1}{1+\kappa} \sum_{a \in \mathcal{A}} \delta_a^p t_a(\overline{\mathbf{v}}) \leq \min_{p' \in \mathcal{P}_w} \sum_{a \in \mathcal{A}} \delta_a^{p'} \overline{\lambda}_a^w t_a(\overline{\mathbf{v}}) \leq \min_{p' \in \mathcal{P}_w} \sum_{a \in \mathcal{A}} \delta_a^{p'} t_a(\overline{\mathbf{v}}),$$

which is equivalent to the following path flow form:

$$c_p(\overline{\mathbf{f}}) \leq (1+\kappa) \min_{p' \in \mathcal{P}_w} c_{p'}(\overline{\mathbf{f}}).$$

Therefore, we conclude that $\overline{\mathbf{f}}$ is a κ-MSatUE traffic flow. □

Proof of Lemma 4.2. We can prove both directions by observing that

$$\begin{aligned}
\sum_{w \in \mathcal{W}} \sum_{p \in \mathcal{P}_w} \widetilde{c}_p^w(\overline{\mathbf{f}})(f_p - \overline{f}_p) &= \sum_{w \in \mathcal{W}} \sum_{p \in \mathcal{P}_w} \sum_{a \in \mathcal{A}} \delta_a^p \lambda_a^w t_a(\overline{\mathbf{v}})(f_p - \overline{f}_p) \\
&= \sum_{w \in \mathcal{W}} \sum_{a \in \mathcal{A}} \lambda_a^w t_a(\overline{\mathbf{v}}) \left(\sum_{p \in \mathcal{P}_w} \delta_a^p f_p - \sum_{p \in \mathcal{P}_w} \delta_a^p \overline{f}_p \right) \\
&= \sum_{w \in \mathcal{W}} \sum_{a \in \mathcal{A}} \lambda_a^w t_a(\overline{\mathbf{v}})(x_a^w - \overline{x}_a^w).
\end{aligned}$$

□

Proof of Lemma 4.5. From Theorem 2 of Christodoulou et al. (2011), we have

$$C(\mathbf{f}^{\kappa}) \leq \zeta(\kappa, n) C(\mathbf{f}) \quad \forall \mathbf{f} \in \mathbf{F}. \tag{A19}$$

Picking $\mathbf{f} = \mathbf{f}^0$ in (A19), we obtain the upper bound on $C(\mathbf{f}^{\kappa})$. Inequalities involving $C(\mathbf{f}^*)$ are from the fact $C(\mathbf{f}^*) \leq C(\mathbf{f})$ for all $\mathbf{f} \in \mathbf{F}$. □

Proof of Lemma 4.6. By simple comparison, we can show

$$\begin{aligned} C((1+\kappa)\mathbf{f}) = Z((1+\kappa)\mathbf{v}) &= \sum_{a\in\mathcal{A}} \left(\sum_{m=0}^{n} b_{am} \Big((1+\kappa)\mathbf{d}_{am}^{\top}\mathbf{v} \Big)^{m} \right) (1+\kappa) v_a \\ &\leq (1+\kappa)^{n+1} \sum_{a\in\mathcal{A}} \left(\sum_{m=0}^{n} b_{am} \Big(\mathbf{d}_{am}^{\top}\mathbf{v} \Big)^{m} \right) v_a \\ &= (1+\kappa)^{n+1} Z(\mathbf{v}) \\ &= (1+\kappa)^{n+1} C(\mathbf{f}) \end{aligned}$$

where $\mathbf{v}$ is the arc flow vector from $\mathbf{f} \mapsto \mathbf{v}$. □

Proof of Lemma 4.7. This is a minor variant to the proof of Englert et al. (2010, Theorem 3). Since $\mathbf{v}^0 \in \mathbf{F}$ and $\widehat{\mathbf{v}}^0 \in \mathbf{F}_{1+\kappa}$ are PRUE flows that minimize $\Phi(\cdot)$ over their corresponding feasible sets, we have

$$\Phi(\mathbf{v}^0) \leq \Phi\Big(\frac{\widehat{\mathbf{v}}^0}{1+\kappa} \Big) \quad \text{and} \quad \Phi(\widehat{\mathbf{v}}^0) \leq \Phi\big((1+\kappa)\mathbf{v}^0 \big),$$

which imply

$$(1+\kappa)^{n+1} \sum_{a\in\mathcal{A}} \frac{b_a}{n+1} (v_a^0)^{n+1} \leq \sum_{a\in\mathcal{A}} \frac{b_a}{n+1} (\widehat{v}_a^0)^{n+1} \tag{A20}$$

and

$$\sum_{a\in\mathcal{A}} \frac{b_a}{n+1} (\widehat{v}_a^0)^{n+1} \leq (1+\kappa)^{n+1} \sum_{a\in\mathcal{A}} \frac{b_a}{n+1} (v_a^0)^{n+1} \tag{A21}$$

respectively.

Let us assume that $C(\widehat{\mathbf{f}}^0) > (1+\kappa)^{n+1} C(\mathbf{f}^0)$, which is equivalent to

$$(1+\kappa)^{n+1} \sum_{a \in \mathcal{A}} b_a (v_a^0)^{n+1} < \sum_{a \in \mathcal{A}} b_a (\widehat{v}_a^0)^{n+1}. \tag{A22}$$

From $n \times$ (A20) $+ ((n+1)(1+\kappa)^n - 1) \times$ (A21) $+ ((1+\kappa)^n - 1) \times$ (A22), we obtain

$$\theta_1 \sum_{a \in \mathcal{A}} \frac{b_a (v_a^0)^{n+1}}{n+1} < \theta_2 \sum_{a \in \mathcal{A}} \frac{b_a (\widehat{v}_a^0)^{n+1}}{n+1} \tag{A23}$$

where

$$\theta_1 = n \cdot \frac{(1+\kappa)^{n+1}}{n+1} - ((n+1)(1+\kappa)^n - 1) \cdot \frac{(1+\kappa)^{n+1}}{n+1} + ((1+\kappa)^n - 1) \cdot (1+\kappa)^{n+1} = 0$$
$$\theta_2 = n \cdot \frac{1}{n+1} - ((n+1)(1+\kappa)^n - 1) \cdot \frac{1}{n+1} + ((1+\kappa)^n - 1) = 0$$

for all $\kappa \geq 0$. Therefore, (A23) leads to $0 < 0$, which is a contradiction. We conclude that $C(\widehat{\mathbf{f}}^0) \leq (1+\kappa)^{n+1} C(\mathbf{f}^0)$. □

Proof of Theorem 4.1. Since $\widehat{\mathbf{v}}^0 \in \mathbf{F}_{1+\kappa}$ is an user equilibrium flow that minimizes $\Phi(\cdot)$, we have

$$\Phi(\widehat{\mathbf{v}}^0) \leq \Phi\big((1+\kappa)\overline{\mathbf{v}}\big),$$

which implies

$$\sum_{a \in \mathcal{A}} \frac{b_a (\widehat{v}_a^0)^{n+1}}{n+1} \leq \sum_{a \in \mathcal{A}} \frac{b_a \big((1+\kappa)\overline{v}_a\big)^{n+1}}{n+1} = (1+\kappa)^{n+1} \sum_{a \in \mathcal{A}} \frac{b_a (\overline{v}_a)^{n+1}}{n+1}. \tag{A24}$$

Since $\overline{\mathbf{v}} \in \mathbf{V}$ is a solution to UE-PE-$\mathbf{V}$, we have

$$\Psi(\overline{\mathbf{v}}; \boldsymbol{\lambda}) \leq \Psi\Big(\frac{\widehat{\mathbf{v}}^0}{1+\kappa}; \boldsymbol{\lambda}\Big),$$

for some $\boldsymbol{\lambda}$. Therefore, we have

$$\sum_{a\in\mathcal{A}} \frac{\lambda_a b_a(\overline{v}_a)^{n+1}}{n+1} \leq \sum_{a\in\mathcal{A}} \frac{\lambda_a b_a(v_a^0)^{n+1}}{(n+1)(1+\kappa)^{n+1}} = \frac{1}{(1+\kappa)^{n+1}} \sum_{a\in\mathcal{A}} \frac{\lambda_a b_a(\widehat{v}_a^0)^{n+1}}{n+1}.$$

Since $\lambda_a \in [\frac{1}{1+\kappa}, 1]$, we obtain

$$\frac{1}{1+\kappa} \sum_{a\in\mathcal{A}} \frac{b_a(\overline{v}_a)^{n+1}}{n+1} \leq \frac{1}{(1+\kappa)^{n+1}} \sum_{a\in\mathcal{A}} \frac{b_a(\widehat{v}_a^0)^{n+1}}{n+1}$$

which implies

$$(1+\kappa)^n \sum_{a\in\mathcal{A}} \frac{b_a(\overline{v}_a)^{n+1}}{n+1} \leq \sum_{a\in\mathcal{A}} \frac{b_a(\widehat{v}_a^0)^{n+1}}{n+1} \tag{A25}$$

Let us assume that $C(\overline{\mathbf{f}}) > C(\widehat{\mathbf{f}}^0)$, which is equivalent to

$$\sum_{a\in\mathcal{A}} b_a(\widehat{v}_a^0)^{n+1} < \sum_{a\in\mathcal{A}} b_a(\overline{v}_a)^{n+1} \tag{A26}$$

From $A \times$ (A24) $+ B \times$ (A25) $+ C \times$ (A26) for any positive constants A, B and C, we obtain

$$\theta_1 \sum_{a\in\mathcal{A}} \frac{b_a(\widehat{v}_a^0)^{n+1}}{n+1} < \theta_2 \sum_{a\in\mathcal{A}} \frac{b_a(\overline{v}_a)^{n+1}}{n+1} \tag{A27}$$

where

$$\begin{aligned} \theta_1 &= A - B + C(n+1) \\ \theta_2 &= A(1+\kappa)^{n+1} - B(1+\kappa)^n + C(n+1). \end{aligned}$$

In particular, consider A, B and C as follows:

$$\begin{aligned} A &= (n+1)\big((1+\kappa)^n - 1\big) \\ B &= (n+1)\big((1+\kappa)^n - 1\big) + (n+1)\kappa(1+\kappa)^{n+1} \\ C &= \kappa(1+\kappa)^{n+1} \end{aligned}$$

We observe that A, B and C are all positive and $\theta_1 = 0$. We also see that

$$\theta_2 = -(n+1)\kappa^2(1+\kappa)^n\big((1+\kappa)^{n+1} - 1\big) \leq 0$$

for all $\kappa \geq 0$ and $n \geq 0$, which leads to a contradiction. Therefore, we have

$$C(\overline{\mathbf{f}}) \leq C(\widehat{\mathbf{f}}^0) \leq (1+\kappa)^{n+1} C(\mathbf{f}^0),$$

where the last inequality is from Lemma 4.7. This completes the proof. □

Proof of Theorem 4.2. By slightly modifying the proof of Theorem 4.3, we can show $C(\mathbf{f}^\kappa) \leq C(\widehat{\mathbf{f}}^0)$. By Lemmas 4.6 and 4.7, we complete the proof. □

Proof of Lemma 4.8. Let $\overline{\mathbf{f}} = (1+\kappa)\mathbf{f}^0$, and $\overline{\mathbf{v}} = (1+\kappa)\mathbf{v}^0$ for the corresponding arc flow vectors. If the condition

$$\sum_{a \in \mathcal{A}} \left(\sum_{m=0}^{n} \lambda_{am} b_{am} \left(\mathbf{d}_{am}^\top \overline{\mathbf{v}} \right)^m \right) (v'_a - \overline{v}_a) \geq 0 \qquad \forall \mathbf{v}' \in \mathbf{V}_{1+\kappa} \tag{A28}$$

holds for some constants $\lambda_{am} \in [\frac{1}{1+\sigma}, 1]$ for $m = 0, 1, ..., n$ and $a \in \mathcal{A}$, then we can find $\lambda_a \in [\frac{1}{1+\sigma}, 1]$ such that

$$\lambda_a \sum_{m=0}^{n} b_{am} \left(\mathbf{d}_{am}^\top \overline{\mathbf{v}} \right)^m = \sum_{m=0}^{n} \lambda_{am} b_{am} \left(\mathbf{d}_{am}^\top \overline{\mathbf{v}} \right)^m$$

for all $a \in \mathcal{A}$; consequently, by Lemmas 4.1 and 4.3, $\overline{\mathbf{f}}$ is a σ-MSatUE flow in $\mathbf{F}_{1+\kappa}$.

Since $\mathbf{v}^0$ is PRUE for $\mathbf{V}$, we know that

$$\sum_{a \in \mathcal{A}} \left(\sum_{m=0}^{n} b_{am} \left(\mathbf{d}_{am}^\top \mathbf{v}^0 \right)^m \right) (v_a - v_a^0) \geq 0 \qquad \forall \mathbf{v} \in \mathbf{V}.$$

Therefore

$$\sum_{a\in\mathcal{A}}\left(\sum_{m=0}^{n}\frac{1}{(1+\kappa)^m}b_{am}\Big((1+\kappa)\mathbf{d}_{am}^{\top}\mathbf{v}^0\Big)^m\right)((1+\kappa)v_a-(1+\kappa)v_a^0)\geq 0 \qquad \forall\mathbf{v}\in\mathbf{V}.$$

Letting for all $a\in\mathcal{A}$

$$\begin{aligned}\lambda_{am}&=\frac{1}{(1+\kappa)^m}, \qquad m=0,1,...,n\\ \overline{v}_a&=(1+\kappa)v_a^0,\\ v_a'&=(1+\kappa)v_a,\end{aligned}$$

we observe that $\lambda_{am}\in[\frac{1}{1+\sigma},1]$ and we obtain (A28); hence proof. □

Proof of Theorem 4.3. We decompose $\mathcal{P}_w$ for each OD pair w into the following four subsets:

$$\begin{aligned}\mathcal{P}_w^1&=\{p\in\mathcal{P}_w:\widehat{f}_p^{\sigma}>0,\ f_p^{\kappa}>0,\ \widehat{f}_p^{\sigma}-f_p^{\kappa}\geq 0\},\\ \mathcal{P}_w^2&=\{p\in\mathcal{P}_w:\widehat{f}_p^{\sigma}>0,\ f_p^{\kappa}>0,\ \widehat{f}_p^{\sigma}-f_p^{\kappa}<0\},\\ \mathcal{P}_w^3&=\{p\in\mathcal{P}_w:\widehat{f}_p^{\sigma}>0,\ f_p^{\kappa}=0\},\\ \mathcal{P}_w^4&=\{p\in\mathcal{P}_w:\widehat{f}_p^{\sigma}=0,\ f_p^{\kappa}>0\}.\end{aligned}$$

We ignore cases with $\widehat{f}_p^{\sigma}=0$ and $f_p^{\kappa}=0$. Note that $\widehat{f}_p^{\sigma}-f_p^{\kappa}>0$ for $p\in\mathcal{P}_w^3$ and $\widehat{f}_p^{\sigma}-f_p^{\kappa}<0$ for $p\in\mathcal{P}_w^4$. From the definition of MSatUE flows, we have

$$\begin{aligned}\widehat{f}_p^{\sigma}>0 &\implies c_p(\widehat{\mathbf{f}}^{\sigma})\leq(1+\sigma)\mu_w(\widehat{\mathbf{f}}^{\sigma}),\\ f_p^{\kappa}>0 &\implies c_p(\mathbf{f}^{\kappa})\leq(1+\kappa)\mu_w(\mathbf{f}^{\kappa}),\end{aligned}$$

for all $p \in \mathcal{P}_w, w \in \mathcal{W}$. In addition, $\mu_w(\widehat{\mathbf{f}}^{\sigma}) \leq c_p(\widehat{\mathbf{f}}^{\sigma})$ and $\mu_w(\mathbf{f}^{\kappa}) \leq c_p(\mathbf{f}^{\kappa})$ for all $p \in \mathcal{P}$ by definition. Therefore, we have

$$
\begin{aligned}
&\sum_{p\in\mathcal{P}}[c_p(\widehat{\mathbf{f}}^{\sigma}) - c_p(\mathbf{f}^{\kappa})](\widehat{f}_p^{\sigma} - f_p^{\kappa}) \\
&\leq \sum_{w\in\mathcal{W}} \Bigg\{ \sum_{p\in\mathcal{P}_w^1} \Big[(1+\sigma)\mu_w(\widehat{\mathbf{f}}^{\sigma}) - \mu_w(\mathbf{f}^{\kappa})\Big](\widehat{f}_p^{\sigma} - f_p^{\kappa}) + \sum_{p\in\mathcal{P}_w^2} \Big[\mu_w(\widehat{\mathbf{f}}^{\sigma}) - (1+\kappa)\mu_w(\mathbf{f}^{\kappa})\Big](\widehat{f}_p^{\sigma} - f_p^{\kappa}) \\
&\qquad + \sum_{p\in\mathcal{P}_w^3} \Big[(1+\sigma)\mu_w(\widehat{\mathbf{f}}^{\sigma}) - \mu_w(\mathbf{f}^{\kappa})\Big](\widehat{f}_p^{\sigma} - f_p^{\kappa}) + \sum_{p\in\mathcal{P}_w^4} \Big[\mu_w(\widehat{\mathbf{f}}^{\sigma}) - (1+\kappa)\mu_w(\mathbf{f}^{\kappa})\Big](\widehat{f}_p^{\sigma} - f_p^{\kappa}) \Bigg\} \\
&= \sum_{w\in\mathcal{W}} \Bigg\{ \sum_{p\in\mathcal{P}_w} \Big[\mu_w(\widehat{\mathbf{f}}^{\sigma}) - \mu_w(\mathbf{f}^{\kappa})\Big](\widehat{f}_p^{\sigma} - f_p^{\kappa}) + \sigma \sum_{p\in\mathcal{P}_w^1\cup\mathcal{P}_w^3} \mu_w(\widehat{\mathbf{f}}^{\sigma})(\widehat{f}_p^{\sigma} - f_p^{\kappa}) \\
&\qquad\qquad - \kappa \sum_{p\in\mathcal{P}_w^2\cup\mathcal{P}_w^4} \mu_w(\mathbf{f}^{\kappa})(\widehat{f}_p^{\sigma} - f_p^{\kappa}) \Bigg\} \\
&\leq \sum_{w\in\mathcal{W}} \Bigg\{ \sum_{p\in\mathcal{P}_w} \Big[\mu_w(\widehat{\mathbf{f}}^{\sigma}) - \mu_w(\mathbf{f}^{\kappa})\Big](\widehat{f}_p^{\sigma} - f_p^{\kappa}) + \sigma \sum_{p\in\mathcal{P}_w} \max\{\mu_w(\widehat{\mathbf{f}}^{\sigma}), \mu_w(\mathbf{f}^{\kappa})\}\Big|\widehat{f}_p^{\sigma} - f_p^{\kappa}\Big| \Bigg\} \\
&\leq \sum_{w\in\mathcal{W}} \sum_{p\in\mathcal{P}_w} \Big[\mu_w(\widehat{\mathbf{f}}^{\sigma}) - \mu_w(\mathbf{f}^{\kappa})\Big](\widehat{f}_p^{\sigma} - f_p^{\kappa}) + \sigma \sum_{p\in\mathcal{P}} \max\{c_p(\widehat{\mathbf{f}}^{\sigma}), c_p(\mathbf{f}^{\kappa})\}\Big|\widehat{f}_p^{\sigma} - f_p^{\kappa}\Big|.
\end{aligned}
$$

From (4.26), we obtain

$$
\begin{aligned}
0 &\leq \sum_{w\in\mathcal{W}} \sum_{p\in\mathcal{P}} \Big[\mu_w(\widehat{\mathbf{f}}^{\sigma}) - \mu_w(\mathbf{f}^{\kappa})\Big](\widehat{f}_p^{\sigma} - f_p^{\kappa}) \\
&= \sum_{w\in\mathcal{W}} \Big[\mu_w(\widehat{\mathbf{f}}^{\sigma}) - \mu_w(\mathbf{f}^{\kappa})\Big]\Bigg(\sum_{p\in\mathcal{P}} \widehat{f}_p^{\sigma} - \sum_{p\in\mathcal{P}} f_p^{\kappa}\Bigg) \\
&= \sum_{w\in\mathcal{W}} \Big[\mu_w(\widehat{\mathbf{f}}^{\sigma}) - \mu_w(\mathbf{f}^{\kappa})\Big](\widehat{Q}_w - Q_w) \\
&= \kappa \sum_{w\in\mathcal{W}} \mu_w(\widehat{\mathbf{f}}^{\sigma})Q_w - \kappa \sum_{w\in\mathcal{W}} \mu_w(\mathbf{f}^{\kappa})Q_w \\
&= \frac{\kappa}{1+\kappa} \sum_{w\in\mathcal{W}} \mu_w(\widehat{\mathbf{f}}^{\sigma})\widehat{Q}_w - \kappa \sum_{w\in\mathcal{W}} \mu_w(\mathbf{f}^{\kappa})Q_w \\
&\leq \frac{\kappa}{1+\kappa} \sum_{w\in\mathcal{W}} \sum_{p\in\mathcal{P}_w} c_p(\widehat{\mathbf{f}}^{\sigma})\widehat{f}_p^{\sigma} - \frac{\kappa}{1+\kappa} \sum_{w\in\mathcal{W}} \sum_{p\in\mathcal{P}_w} c_p(\mathbf{f}^{\kappa})f_p^{\kappa} \\
&= \frac{\kappa}{1+\kappa} C(\widehat{\mathbf{f}}^{\sigma}) - \frac{\kappa}{1+\kappa} C(\mathbf{f}^{\kappa}).
\end{aligned}
$$

Lemmas 4.6 and 4.8 complete the proof. □

Appendix E: Assymetric Nine-node Network

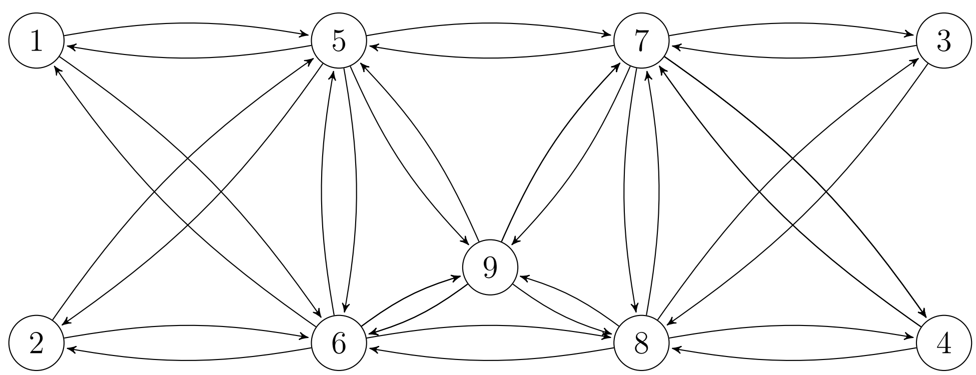


Figure A.1 – Asymmetric nine-node network

In order to test the performance of UE-PE-**X** model in an asymmetric network, we create an asymmetric version of the nine-node network considered by Hearn and Ramana (1998). In the asymmetric nine-node network, which has been shown in Figure A.1, we add a few additional arcs and assume that the arc travel cost function is:

$$t_a(\mathbf{v}) = A_a + B_a \left(\frac{0.5v_{\widehat{a}} + v_a}{C_a} \right)^4 \tag{A29}$$

where $\widehat{a}$ is the flow in the opposite arc. Thus, the arc travel function depends not only on the flow in that arc, but also on the flow in the arc in opposite direction. The values of parameters A_a, B_a and C_a are given in Table A.1 for each arc.

Table A.1 – Asymmetric nine-node network link cost function parameters

a	A_a	B_a	C_a
(1,5)	12	1.80	5
(1,6)	18	2.70	6
(2,5)	35	5.25	3
(2,6)	35	5.25	9
(5,6)	20	3.00	9
(5,7)	11	1.65	2
(5,9)	26	3.90	8
(6,8)	33	4.95	6
(6,9)	30	4.50	8
(7,3)	25	3.75	3
(7,4)	24	3.60	6
(7,8)	19	2.85	2
(8,3)	39	5.85	8
(8,4)	43	6.45	6
(9,7)	26	3.90	4
(9,8)	30	4.50	8
(5,1)	12	1.80	5
(6,1)	18	2.70	6
(5,2)	35	5.25	3
(6,2)	35	5.25	9
(6,5)	20	3.00	9
(7,5)	11	1.65	2
(9,5)	26	3.90	8
(8,6)	33	4.95	6
(9,6)	30	4.50	8
(3,7)	25	3.75	3
(4,7)	24	3.60	6
(8,7)	19	2.85	2
(3,8)	39	5.85	8
(4,8)	43	6.45	6
(7,9)	26	3.90	4
(8,9)	30	4.50	8