Planning and Operation of Automated Taxi Systems

Liang, Xiao

DOI
10.4233/uuid:ff4e3db6-69e5-4b8f-be68-a94c2d84b8bb

Publication date
2019

Document Version
Final published version

Citation (APA)
Liang, X. (2019). *Planning and Operation of Automated Taxi Systems*. [Dissertation (TU Delft), Delft University of Technology]. TRAIL Research School. https://doi.org/10.4233/uuid:ff4e3db6-69e5-4b8f-be68-a94c2d84b8bb

Planning and Operation of Automated Taxi Systems

Xiao Liang

Delft University of Technology

This thesis is the result of a PhD research partly funded by Delft University of Technology (TU Delft), China Scholarship Council (CSC) under the grant 201406260192 and Dutch rail infrastructure manager ProRail under the project D2D100%EV.

Cover illustration: Qu Hu

Planning and Operation of Automated Taxi Systems

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus, Prof.dr.ir. T.H.J.J. van den Hagen,
chair of the Board for Doctorates
to be defended publicly on
Monday 30, September 2019 at 12:30 o'clock

by

Xiao LIANG
Master of Communication and Transportation Engineering,
Tongji University, China
born in Liaoning, China.

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus	chairperson
Prof.dr.ir. B. van Arem	Delft University of Technology, promotor
Dr.ir. G.H.A. Correia	Delft University of Technology, copromotor

Independent members:

Prof.dr.ir. K.I. Aardal	Delft University of Technology
Prof.dr.ir. L.A. Tavasszy	Delft University of Technology
Prof.dr.ir. A.P. Antunes	University of Coimbra
Dr. ir. N.A.H. Agatz	Erasmus University Rotterdam

Other members:

Dr.ir. K. An	Monash University

TRAIL Thesis Series no. T2019/13, the Netherlands Research School TRAIL

TRAIL
P.O. Box 5017
2600 GA Delft
The Netherlands
E-mail: info@rsTRAIL.nl

ISBN: 978-90-5584-255-1

Printed in the Netherlands

Dedicated to my husband, Qu Hu.

Preface

It is with mixed feelings that I look back now that the journey of my PhD research comes to an end. After finishing this precious journey, with all these stressful, anxious, and of course, delightful memories, I am really happy and proud that I never let any difficulty stop me. I also sincerely appreciate everyone who has helped me during the past years.

First of all, I would like to thank my promotor, Prof. Bart van Arem, not only because of the opportunity he gave me to start my PhD study in the Department Transport & Planning at Delft University of Technology but also because of the support he provided during these years. Although he has a very busy schedule, he always paid close attention to my research. I really appreciate these progress meetings and valuable suggestions, which made sure I was always working in the right direction.

Secondly, I would like to thank my daily supervisor, Dr Gonçalo Correia. Gonçalo is the one who helped me get started with optimization research and guided me through my whole PhD. He taught me a mass of knowledge and experience for my research at the starting stage of my PhD. He also gave me the flexibility to explore the new directions of my research when I was more experienced. I really enjoyed all the discussions with him. He always provided sharp ideas and gave clear and understandable explanations. I really appreciate working with him, which sharpened my skills as a researcher. Moreover, I would like to thank Dr Kun An for her thoughtful and thorough guidance during my visit to Monash University and the following collaboration research. Kun showed me how to be an outstanding non-English-speaking female researcher. I can always learn this from her.

Thirdly, I would like to thank all the friends and colleagues during these years. I am very grateful that I have been working with a group of excellent colleagues: Haneen, Oded, Fei, Pengling, Nikola, Yongqiu, Xavier, Yan, Silvia, Na, Yihong, Vincent, Ding, Senlei and

Rongqin. I also thank Dr Meng Wang and Dr Yufei Yuan, for their help and support for the past years. Thanks to all my office mates, Hugo, Han, Freddy, Kai, Lígia, Meiqi, Peyman, Malvika, Shi and Wenjing, for the nice conversations and helps. I also would like to thank Dehlaila, Priscilla, Conchita and Moreen for their kindness and help for my study and defence.

I wish to give thanks to my lunch mates and hiking mates, Dr Shijie Li, Dr Jialun Liu and Wenhua Qu, for the pleasant time we spent together. I really appreciate the help of Dr Mo Zhang, who is a perfect friend, colleague, and neighbour. I also thank Dr Lan Song and Yaqing Shu for their countless warm hospitalities with great food and board games. I would like to give special thanks to Lin Xiao, who has been the best friend, sister, colleague, schoolmate and neighbour of my life. Words cannot express how lucky I felt for the past eight years since we knew each other. I also feel lucky that you will keep being my best friend, sister, neighbour (I wish) for the following years. I am grateful to Lin Xiao and Dr Zhen Yang for their kind help since we move to the Netherlands. And I will give special thanks to three lovely babies, Jin Yang, Yuyao Shu and Yufeng Shu for the warm and beautiful time we spent together.

I appreciate that I have the best parents in the world. My mom and dad are straightforward, but I know that they love me more than anyone in the world. My mom teaches me how to balance well between being a good female worker and a good wife. My dad teaches me how to be a good life planner with everything optimized. I am grateful for their unconditionally emotional and financial support.

I would like to thank my husband Qu Hu. Thank you for all the supports and discussions during my bachelor, master and PhD journey. I really appreciate all the sweet moments we have ever had.

Xiao Liang

Delft, June 2019

Contents

List of Figures

List of Tables

List of Abbreviations

AT	Automated taxi
AV	Automated vehicle
BPR	Bureau of Public Roads
CV	Coefficient of variation
DARP	Dial-a-ride-problem
IP	Integer programming
LIP	Linear integer programming
NLIP	non-linear integer programming
NP-hardness	non-deterministic polynomial-time hardness
TNC	transportation network companies
VRP	vehicle routing problem

Chapter 1

Introduction

This chapter introduces the topic of this thesis by presenting the current knowledge and the main challenges in the planning and operation of automated taxis. It is structured as follows. Section 1.1 introduces the research background of automated vehicles used as taxis. Section 1.2 proposes the research questions that should be answered in the thesis based on the literature review. Section 1.3 describes the research approach which is followed by the research scope in Section 1.4. Section 1.5 discusses the scientific and practical contributions of this thesis. Finally, the outline of the thesis is presented in Section 1.6.

1.1 Background of Automated Taxis

An automated vehicle (AV), also known as a driverless car and a self-driving car is an advanced type of vehicle that can drive itself on existing roads. The SAE (Society of Automotive Engineers) identifies six levels of vehicle automation from level 0 (no automation) to level 5 (full automation), where the driving automation system has different levels of responsibility for controlling a car. Within these six levels, the highest level, full driving automation, can be defined as the "sustained and unconditional performance by a driving automation system of the entire dynamic driving task without any expectation that a user will respond to a request to intervene" (SAE International, 2018). This means that vehicles in fully automated mode are not only able to monitor the driving environment and execute the dynamic driving tasks (e.g. steering, braking, responding to events, determining when to change lanes), but they are also capable to do so in all driving environments (e.g. expressway merging, high-speed cruising, low-speed traffic congestion). On the other hand, fully automated driving mode changes human's role when engaging the driving automation: the one who used to be the driver in the conventional vehicle becomes just a passenger like the other occupants. Therefore, this new technology has the potential to remarkably change the way in which motorized transport systems operate.

According to some experts, driving automation is expected to bring significant benefits such as higher safety, lower traffic congestion, lower transport costs, lower usage of parking space, etc. Once vehicle automation technology has been fully developed, traffic collisions caused by human factors should be significantly reduced, since the driver's error contributes to more than 90% of the traffic accidents (Fagnant and Kockelman, 2014; KPMG, 2012). AVs could have higher speed limits, smoother rides and lead to higher road capacity due to the reduction of the headway required (Hoogendoorn et al., 2014). In addition, AVs could reduce labour costs and relieve travellers from driving to other activities like leisure or work (Zhang, 2014). The insurance costs could also be reduced due to safer driving (Light, 2012). Widespread use of AVs can potentially lead to higher traffic efficiency, due to increased traffic flow stability (Anderson et al., 2016; Pyper, 2015; Wang, 2015). With respect to the parking space, if AVs are used as taxis they could satisfy multiple trips continuously throughout a day which could dramatically reduce the demand for parking facilities since human-driven vehicles are reported to be parked for around 95% of the time (Barter, 2013; Bates and Leibling, 2012).

In recent years, most of the research effort has been invested in the technology challenges of creating different automation levels of AVs. Technology companies, public transport companies and vehicle manufacturers have run several AV pilot projects. Waymo, which has originated from the Google self-driving car project and became a subsidiary company since 2016, has cars which have driven more than 10 million miles on public roads in automated mode since 2009. The company plans to implement AVs in different application areas of transportation systems, "from ride-hailing and logistics to public transport and personal vehicles" (Waymo, 2018). By 2017, Mercedes has vastly expanded its automated driving features on their cars. In addition to the standard features such as an active brake assist,

Mercedes now includes a steering pilot, a parking pilot, a cross-traffic assist system, night–vision cameras with automated danger warnings and braking assist, and various other automated-driving features (Daimler, 2017; Mercedes Blog-Team, 2017; Ward, 2014). Tesla Autopilot is an advanced driver-assistance system offered by Tesla, which includes lane centring, adaptive cruise control, self-parking, etc. (The Tesla Team, 2018). It was first offered in 2014, followed by several versions of hardware and software updates. Based on this system, Tesla intends to provide full self-driving cars in the market in the near future, after millions of miles of real-world driving for calibration purpose (Golson and Bohn, 2016). Uber started to test their self-driving car in Pittsburgh, the USA, since 2016. This real-world testing focuses on improving the technology and ensuring its safety for all road users: pedestrians, cyclists and other drivers. In 2018, Toyota started to work with Uber on AVs. They expect to produce automated minivans by 2021, which combine Toyota's car-making expertise with Uber's AV technique and ride-hailing platform (MARSHALL, 2018). However, technology is not the only challenge involved in the development of AVs. Problems such as safety, liability, legal framework, government regulations, privacy and security concerns, as well as sustainability impacts and transport system control need to be solved before the implementation of AVs in the real-world market.

AVs are predicted to be increasingly used in the future. Bierstede et al. foresee that 25% of the vehicles on the road will be automatic by 2035 and 50% in 2035-2050 (Bierstadt et al., 2014). If the government authorization or personal subscription of ownership prevails, then the AV share may be beyond 75% by 2035 and beyond 95% by 2040. The Victoria Transport Policy Institute forecasts the AV implementation rates based on the assumption that fully AVs are available for sale and legal to drive on public roads around 2020. They predict that by 2050 the AVs will achieve a 40%-60% share of the vehicle fleet and 50%-80% of vehicle travel (Litman, 2018). The previous forecasting approaches are mainly qualitative, while Nieuwenhuijsen et al. (2018) proposed a system dynamics model to study the diffusion of AVs in the future. They conclude that the market penetration of AV is highly uncertain and varies greatly with different levels of technology development and policies adopted. In a pessimistic scenario, strong adoption of automation level 2 (between no automation and full automation) can be achieved by 2025 (51% share) and level 3 by 2100 (64% share). However, in a very optimistic scenario, the market penetration of AVs would be 35% of level 5 by 2025 and 99% of level 5 by 2100. Based on the above, in the medium-run, we should be able to see AVs travelling on the road to undertake transport tasks instead of only for field testing.

A possible area of application for AVs is public transport (Correia et al., 2019; Lamotte et al., 2017). Introducing AVs in public transport may bring benefits to the urban transportation system in two ways. Firstly, vehicle automation could decrease labour needs and improve labour productivity for the public transport industries. The absence of human drivers will directly benefit the public transport operators due to cost reduction which enables them to re-formulate the cost structure. In countries like Japan, public transport companies have suffered from a shortage of bus and taxi drivers for a long time, and this may get worse in the future due to the ageing population (TCA, 2018). However, AVs make it possible to provide transport service when labour is deficient. Secondly, AVs could improve the service quality

of the public transport system for the citizens. For instance, it might shorten the waiting times for taxis and enhance the frequency of mass transport systems like bus, metro and train. Decreased operation costs might lead to cheaper fares for public transport. Throughout the world, metro, train, tram or bus systems are frequently used as public transport. Unfortunately, inflexibility, long travel time, poor access and egress and insufficient service coverage of these transport systems cause their lower usage in urban, suburban and rural areas (Chong et al., 2011). This kind of transport usually has a centralized management system which uses Intelligent Transport Systems (ITS) technologies for optimal operation of the service. Conversely, taxis are a more convenient mode due to their fast, door-to-door connection, privacy, comfort, 24-hour service and lack of parking fees (Salanova et al., 2011). This high-quality service results in a higher price for using a taxi, especially compared to the mass public transport. With the emergence of AVs, there could be a reduction in the price of taxis thus making them more competitive with other modes (Krueger et al., 2016).

The concept of automated taxis (ATs) is supposed to offer a seamless door-to-door service within a city area for all passengers. Some researchers have given attention to testing the effect of using ATs on urban transport, especially looking at mixed traffic combination with conventional vehicles. Two methods have been widely used to test these impacts: 1) agent-based simulation; 2) mathematical optimization. Martinez and Viegas (2017) used agent-based simulation to build a model to test the introduction of 100% automated fleets of taxis to satisfy transport demand in a city. Results showed that with the subway still in operation each AV could remove 9 out of 10 cars in the city if a maximum 5 min waiting time is to be guaranteed, whilst without metro, the number is reduced to 5 vehicles removed per AV. Fagnant and Kockelman (2014) used a similar method to study the implications of shared ATs and compared them to conventional vehicle ownership and use. Their results indicate that each shared AV could replace around 11 conventional vehicles, but they add up to 10% more travel distance. Spieser et al. (2014) used an analytical mathematical formulation to estimate the number of shared AVs to replace all modes of personal transportation in the case-study city of Singapore. Using the data from a real case study, they were able to conclude that a shared-vehicle mobility solution could meet the personal mobility needs of the entire population with 1/3 of the number of passenger vehicles currently in operation. Madadi et al. (2019) explore the travel impacts of vehicle automation level 3-4 in an urban road network with mixed traffic, i.e. automated driving is only allowed on some selected roads and for the remaining road is with manual driving. The results indicate that with higher adoption of high-level automation, total travel costs and total travel time are decreased, while total travel distance is slightly increased. Based on these studies cited above, it is possible to conclude that ATs would be more beneficial and efficient in urban transport and could potentially replace many conventional vehicles while providing the same level of system performance.

With the advent of automation, using ATs in urban transport systems creates a new type of shared economy, which is similar to traditional carsharing. Traditional carsharing systems provide more sustainable urban mobility compared to private cars (Shaheen et al., 1999). Vehicles in these systems have higher utilization rates when compared to the privately-owned

ones (Celsor and Millard-Ball, 2007; Jorge et al., 2015b, 2015a; B. Li et al., 2016; Ma et al., 2017; Schuster et al., 2005). However, the shared-use vehicles must be relocated between different areas due to the demand imbalance, which leads to time and monetary costs. Moreover, traditional carsharing systems usually have either fixed vehicle stations for location-based systems or random parking locations for free-floating systems. Hence the users must walk to reach the vehicles. Using AVs in a carsharing system could reduce vehicles' relocation costs and eliminate users' self-serve access to the vehicles. Therefore, shared ATs are expected to be as flexible and convenient as traditional taxis and as sustainable and economical as carsharing. In the near future, we may be able to see the situation in which hundreds or even thousands of AVs will be on the road replacing private vehicles accounting for the majority of people's daily trips.

Ride-sharing is another important component in shared mobility, which aims to bring together travellers who have similar itineraries and time schedules to share rides (Agatz et al., 2012, 2011; Correia and Viegas, 2011; Krueger et al., 2016; Long et al., 2018; Mahmoudi and Zhou, 2016; Masoud and Jayakrishnan, 2017; Najmi et al., 2017; Schaller, 2018; Sethi et al., 1991). The large demand and the low occupancies in private transport in peak hours create traffic congestion in many urban areas. Ride-sharing allows people to use transport capacity more efficiently (Furuhata et al., 2013). Currently, ride-sharing is happening for example with Uber-pool systems whereby a person may request a ride at a lower price but be willing to share with other passengers. The transport service can be provided in two ways: by a private car owner or a transport company. In the first way, users can provide a ride as a driver or ask for a ride as a passenger. Once the travel requests are submitted, there will be matching between the drivers and the riders. In the matching process, the key constraint is the time schedules of the rides. The drivers should have sufficient time flexibility since they need to accomplish the pick-up and drop-off of the passengers and then arrive at their own destinations. If ATs are used in the service scheme of ride-sharing, they will provide the opportunity to transform the role of the drivers into passengers, who have no need to stay in the vehicles for the whole ride. In the second ride-sharing system, a company-hired driver, who does not have the mobility demand, undertakes the driving tasks for all the ride-sharing passengers. When the driver is replaced by the automated system, this person is not needed any more, which saves the transport company employment costs and vehicles' carrying capacity. In summary, with the benefits gained by using vehicle automation, ATs with ride-sharing have the potential to increase the transport capacity and improve the service quality of the public transport system.

1.2 Problem statement and research questions

With automation technology maturing, we may be able to see the situation in which hundreds or even thousands of ATs will be on the road replacing private vehicles accounting for the majority of people's daily trips. Existing research has demonstrated that people could benefit a lot from using vehicle automation in urban mobility. However, little attention has been

devoted to the usage of a fleet of AVs as taxis and their effect on a real-scale road network. We need to think on how automated driving can serve mobility and what is the best way to introduce this technology as part of the existing transport networks: either as a complementary mode or as a substitution mode. But the scientific evidence is still lacking. Some existing research pays attention to this kind of problems (e.g. location problem, vehicle routing problem, traffic assignment problem, etc.) within a conventional vehicle system. However, it is an emerging area to study the system design based on the features of AVs, which is the research gap this thesis is going to fill. In this research, we consider ATs and aim at identifying their role in urban mobility systems, looking at the problem from a taxi company's perspective. To be more specific, the objective of this thesis is to contribute to the planning and operational strategies that these AT systems should follow in order to satisfy urban mobility demand.

To achieve efficient planning and operation for ATs in the urban mobility system, the main research question proposed in this thesis is:

How should an AT system be designed in order to optimally serve people's urban travel demand?

To address this main question, the following main sub-research questions will be answered:

1. *How should the service area of ATs be designed for the last mile service?*
2. *How should the ATs choose the routes considering traffic congestion?*
3. *How should people's real-time demand to be satisfied with an AT system?*
4. *How should the route choice problem with ATs be solved in an efficient way?*

The first sub-research question focuses on defining the service area of ATs according to the travel demand distribution in time and space during the planning stage of adopting an AT system. The service area consists of several geographical zones where people can start and end their trips. The second sub-research question focuses on optimizing the route choice when ATs are assigned to pick-up and deliver passengers during the operation stage. The third sub-research question is also for the operation stage, which focuses on establishing a time-dimension framework to address the real-time demand of ATs. This makes ATs' route choice results flexible enough to handle the new customers being generated through time. The fourth sub-research question focuses on the mathematical challenge of solving the problems proposed in sub-research question 2 and 3.

1.3 Research scope

This thesis focuses on the AV's application in passenger transport, using AVs as taxis in an urban area. An AT system is designed to offer seamless door-to-door service within a city area, meaning that the intercity transport demand is not considered.

We consider two kinds of services in this thesis: last-mile and full coverage of an urban area. The last mile transport service is for accessing and egressing from train stations to the service

zones. This means that the requests to use the system between two service zones are not considered. Then we relax this assumption and open the AT service to the full coverage of the city demand. Therefore, the passengers can be delivered between any pair of nodes within the city area.

The travel time is considered static and dynamic in different parts in this thesis. Firstly, ATs are assumed to be mixed with conventional vehicles with deterministic travel times. Then we assume that a large number of ATs are used to meet people's travel demand replacing all private vehicles, meaning that the congestion caused by the ATs themselves cannot be ignored. Therefore, we establish the flow based travel time function and include them in the following optimization model.

The optimization models take into account both static and dynamic demand information. Firstly, the optimization process is static, requiring all requests to be booked ahead. Considering the information availability in real-world, we convert the static model to a dynamic one using a rolling-horizon framework. Under this framework, the system is able to address the real-time demand revealed over the day and update the system status periodically.

The AT service is assumed to be serving individual demand first but later ride-sharing is also taken into account. We firstly allow ATs to satisfy individual trips meaning that one AT can only serve one passenger at a time. Then ride-sharing is considered meaning that the passengers can share a ride with others who have similar itineraries and time schedules when they are served by ATs.

1.4 Research approach

Mathematical optimization and computer simulation are two methods widely used by researchers to investigate the effects of using ATs on urban transport. A mathematical optimization problem consists of maximizing or minimizing a function by systematically selecting some input values within a defined domain. It aims to find the best available values of the objective function and the corresponding values of the problem input. Computer simulation is an experiment to imitate the behaviour and the outcomes of a proposed mathematical model. It allows testing the reliability and viability of the chosen model and performing uncertainty and randomness in a real system.

This thesis uses mathematical optimization to answer the above research questions. The purpose of this thesis is to provide a tool to support the decision-making processes both for long-term planning strategies and short-term tactical operations when ATs are going to be applied in the urban transport system. Using optimization is a scientific approach to produce high-quality solutions of the system input and achieve the best value of an objective. When the proposed mathematical optimization model is too complicated that no known polynomial algorithms can solve it in polynomial time, then the model turns into so-called NP-hardness (non-deterministic polynomial-time hardness). Hence, the scientific challenge will also include the solving procedure of the NP-hard optimization models.

For each sub-research question, a method involving a mathematical model is proposed for answering the question. We take into account costs and revenues for the planning and operation of the AT system and set the profit maximization as the objective of the optimization. The models are applied to case-studies to demonstrate their performance and conclude on their applicability.

Firstly, to define the service area of an AT system for the last mile problem, a facility location problem is formulated as an integer programming (IP) model (Huang et al., 2018; Jorge et al., 2012; Juster and Schonfeld, 2013; X. Li et al., 2016). By integrating the trip selection procedure at an operational level, this model is able to select the service zones among all the potential zones, according to the profit maximization objective. Different service schemes are proposed and numerical experiments are applied. This answers research question 1.

Secondly, we define the route choice of ATs and their assignment to clients as a (DARP), which is a variation from the classical vehicle routing problem (VRP). A VRP aims to design the best routes to provide services from a depot to some customers distributed in the network (Laporte, 2009). When the problem involves transporting people from their origins to their destinations with request time windows, it becomes a DARP (Ho et al., 2018). With dynamic travel time, AT's DARP is developed to design the best routes to transport people from their origins to their destinations, according to their desired time window. This formulation involves congested assignment for the AT fleet to decide the route choice, by integrating a non-linear flow based travel time function. A linearized framework is proposed to simplify the optimization model and make the problem solvable. This answers research question 2.

Thirdly, a rolling-horizon framework is proposed to address real-time requests, which is called a dynamic DARP problem. It is a periodic re-optimization which returns to the solving procedure each time an update of the demand occurs. This update is defined as a period of deterministic time in the rolling-horizon framework. Since the planning horizon is divided into multiple smaller periods, it is possible to reduce the scale of the problem, which is also a way to handle the NP-hardness of the problem although abdicating from finding a global optimum. This answers research question 3.

Fourthly, in order to accelerate the solving procedure, a customized Lagrangian relaxation solution algorithm is proposed, which enables to approach the optimal solution. The relaxed model is further decomposed into two sub-problems. The non-linear flow based travel time function remains and the sub-problem related to this is solved by an iterative assignment process. Moreover, ride-sharing is further considered to increase the transport efficiency of the AT system, meaning that the constraints for ATs' seating capacity are added in the mathematical formulation. This answers research question 2, 3 and 4 together

1.5 Main contributions

This thesis focuses on AT system optimization. Mathematical models for designing an AT system and evaluating its performance are developed, and application issues have also been

considered through case-studies. In this section, we highlight the main contributions of this thesis, grouped as scientific contributions and practical contributions.

1.5.1 Scientific contributions

- *Formulating an optimization model for designing the service area of an AT system for last mile accessibility to train stations.*

This fills the gap of providing a model that optimally selects the opening zones for ATs to provide access to train stations. By considering two different service schemes, this model furthermore enables a direct comparison of the service performance between free service (trip reservations are accepted or rejected by the operator according to the profit maximization) and full service (any reservation on a selected zone by the model must be satisfied). Additionally, the model considers the vehicles to be electric thus charging constraints are included in order to guarantee that on average ATs are idle enough time for the charging to happen.

- *Considering the effect of traffic congestion for integrating route choice into AT's DARP.*

These models represent one of the first attempts to use an optimization method to choose ATs' routes while involving traffic congestion since for a growing number of ATs circulating in the city their number will lead to delays. With such consideration, this model is able to reflect the congestion effect on travel time and allow the ATs to choose the optimal routes.

- *Developing static and dynamic frameworks to address AT's DARP.*

This is the first time that AT's DARP is used in two situations with different information availability: static and dynamic. Chapter 3 assumes the demand is deterministic and pre-known during the optimization period, which is defined as a static DARP. Chapter 4 and 5 propose a rolling-horizon framework to deal with the dynamic DARP with real-time demand generating during the optimization period. This framework divides a typical day into several time-horizons and periodically re-optimizes the problem to update the implemented results.

- *Presenting effective solution algorithms for solving AT's DARP.*

This is one of the first studies to address the challenges in solving ATs' DARP problem. This thesis in Chapter 2, 3 and 4 uses commercial software to solve the proposed optimization models. When more realism is included in the problems we want to solve, the models become more complicated and computationally challenging. Therefore, Chapter 5 proposes a solution approach based on a customized Lagrangian relaxation algorithm, which enables to identify a near optimal solution for AT's DARP.

1.5.2 Practical contributions

- *Providing a practical tool for industry and government to design the service area and define the fleet size for ATs at the planning level.*

This thesis in its Chapter 2 applies the service area model to a real city case study which enables industry (including vehicle manufacturers, service providers, etc.) and government to

do the planning for the implementation of the AT system, i.e. selecting the service area and deciding the fleet size.

- *Demonstrating the benefits of using AVs as taxis in terms of costs and service quality for policy-makers.*

This thesis in Chapter 2 compares the performance of an AT system with a human-driven taxi system. Under the same travel demand of citizens, these two systems present a significant difference in terms of operating profit and demand satisfaction rate. This comparison provides the policy-makers with a quantitative result of the benefits of introducing AVs in the public transport system and using them as shared taxis.

- *Providing a practical tool for AT companies to select the routing of ATs at the operational level.*

This thesis in its Chapter 3,4 and 5 applies AT's DARP to a real city case study which enables taxi companies to operate a fleet of ATs by assigning individual passengers to vehicles and selecting the optimal routes between each origin-destination pair. The rolling horizon framework also allows the AT company to address the real-time requests, which has more flexibility to deal with the real-world demand.

- *Revealing the performance of the AT service and the profitability of the AT system for stakeholders.*

The thesis estimates the profit of the AT system under different scenarios, which enables the stakeholders to evaluate the service performance (service coverage, waiting time, delay time, etc.) and the system profitability of implementing ATs. Research findings can also be used to make general recommendations for AT's planning and operation, relating to, for instance, how to guarantee high-quality service, which control parameters will influence the profit.

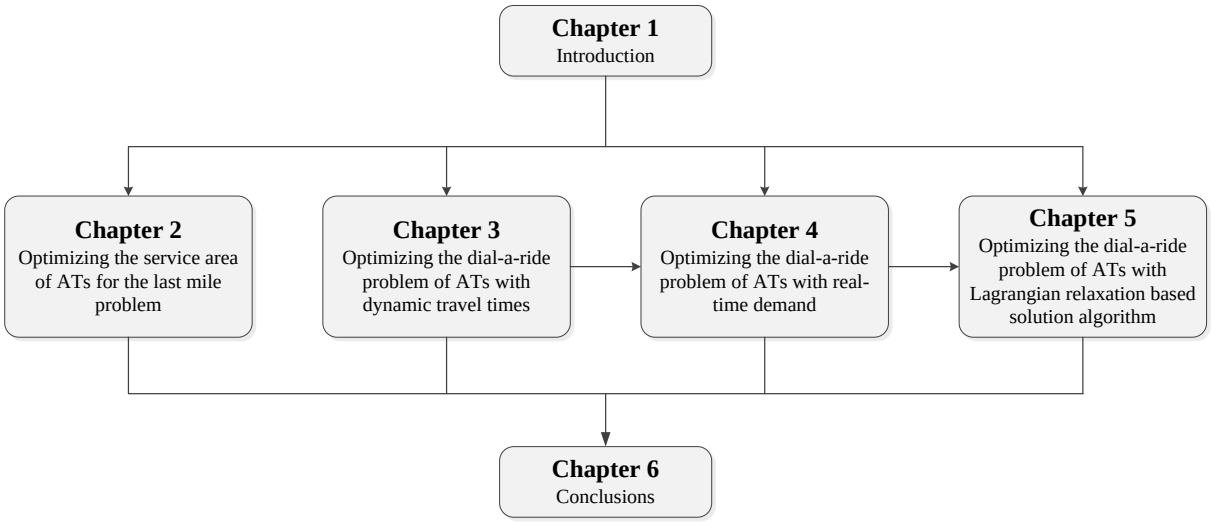


Figure 1.1 Overview of thesis structure

1.6 Outline

Figure 1.1 gives an overview of the structure of the thesis. This thesis consists of 6 chapters, and chapters 2 to 5 are mainly composed of published or under review scientific journal papers. The main research question to be addressed is a complex one, which needs to be simplified in different aspects. This thesis goes through the research elements from a simple setting to a more complex one. The closer one gets to more complexity, the closer it gets to the challenges of a real application. Chapter 3 proposes an IP model to define the routing of the vehicles according to a profit maximization function while depending on dynamic travel times which vary with the flow of the ATs. The model is linearized by transforming a non-linear flow based travel time function into a set of discrete points. This model is applied to a small toy network and the results allow assessing the impact of the AT's movements on traffic congestion and the profitability of the system.

Table 1.1 gives an overview of the elements involved in each chapter of this thesis.

The chapters are organized as follows: Chapter 2 proposes a mathematical model to optimize the service area and trip selection of a transport system with ATs which serves the last mile connection of train trips requiring pre-booking. We apply this model to a case study city Delft, the Netherlands, with its secondary train station Delft Zuid. With the reservation requests being known a priori, the model produces the satisfied demand and the number and location of service zones that lead to a more profitable system.

Chapter 3 proposes an IP model to define the routing of the vehicles according to a profit maximization function while depending on dynamic travel times which vary with the flow of the ATs. The model is linearized by transforming a non-linear flow based travel time function into a set of discrete points. This model is applied to a small toy network and the results allow assessing the impact of the AT's movements on traffic congestion and the profitability of the system.

Table 1.1 Research elements

	AT service scope	Travel time type	Demand informing	Sharing type	Solving algorithm
Chapter 2	Last mile	static travel time	pre-known	Individual trips	commercial solver
Chapter 3	Full coverage	dynamic travel time	pre-known	Individual trips	commercial solver
Chapter 4	Full coverage	dynamic travel time	real-time	Individual trips	commercial solver
Chapter 5	Full coverage	dynamic travel time	real-time	Ride-sharing trips	Lagrangian relaxation

Chapter 4 proposes a rolling horizon framework to divide one day into several periods in which both the real-time and the booked demand will be considered together to determine

ATs' route choice. The application to the city of Delft, the Netherlands demonstrates that the taking into account the effect of the vehicle flows on travel time leads to changes in the system profit, the satisfied percentage of the ATs.

Chapter 5 proposes a non-linear model which is an extension of the one in chapter 3 and 4. This model studies the DARP of ride-sharing ATs in an urban road network, considering dynamic travel times. In this chapter, we design a solution approach based on a customized Lagrangian relaxation algorithm which enables to identify a near optimal solution faster for the proposed model. Numerical experiments for the city of Delft, the Netherlands, are used to demonstrate the solution quality of the proposed algorithm as well as obtaining insights about the AT system performance.

Chapter 2

Optimizing the service area of automated taxis for the last mile problem

The previous chapter introduces the topic of this thesis and reviewed the application of automated vehicles in public transport. When using automated vehicles as taxis, a planning strategy is needed to define the service area of this system. In this chapter, we propose a mathematical model to optimize the service area and trip selection of a transport system with ATs which serves the last mile connection of train trips requiring pre-booking. We consider a city divided into several areas, each of which can be seen as a potential service zone. The service provided by this system includes taking passengers who have finished their train trips to the final destination (egress) or to the train station from the origin (access). We consider the vehicles to be electric which means that charging time has to be taken into consideration as a constraint. At the same time, because taxis are automated, it is possible for them to relocate without a driver. Two models are formulated as IP problems: in the first (model S1), trip reservations are accepted or rejected freely by the operator regardless having their origin/destination on a selected zone; in the second (model S2), any reservation on a selected zone by the model must be satisfied by the system. The model was applied to the Delft Zuid train station in The Netherlands.

The chapter is structured as follows: Section 2.1 reviews the existing research about ATs and last mile problem. Section 2.2 introduces the mathematical models for two different trip selection schemes. Section 2.3 and section 2.4 apply the models to the case study of the Delft

Zuid train station in the city of Delft (the Netherlands) and show the optimal results from the case study. Finally, a discussion on the results and main conclusions drawn from the model application are presented in section 2.5.

This chapter is an edited version of the following paper:

Liang, X., Correia, G.H. de A., van Arem, B., 2016. Optimizing the service area and trip selection of an electric automated taxi system used for the last mile of train trips. *Transportation Research Part E: Logistics and Transportation Review* 93, 115–129. doi:10.1016/j.tre.2016.05.006

2.1 Introduction

Within the last decade, technology development has accelerated the process of vehicle automation. An AV, also known as a driverless car and a self-driving car is an advanced type of vehicle which can drive itself on existing roads and can navigate many types of roadways and environmental contexts with reduced direct human input (Fagnant and Kockelman, 2013). Fully AVs are expected to bring significant benefits, such as mitigating traffic congestion, reducing car crashes, improving fuel efficiency and alleviation the negative impacts on the environment (Bierstadt et al., 2014). Although further evidence is still needed to assess if those advantages are indeed real.

In this work, we do not study the full substitution of traditional transit networks but propose instead to analyse the potential of using ATs as the last mile connection of train trips. Given parking space availability, a properly functioning road infrastructure and smooth traffic, the use of the private automobile is highly attractive especially at longer distances (Ford, 2012). Moreover, in multimodal trips, it has been shown that a relatively high disutility is caused by the access and egress modes of transport (Hess, 2009; Hoogendoorn-Lanser et al., 2006). At the same time, to make transport more efficient, concentrating passengers in higher capacity vehicles such as trains leads to cost and pollution savings, hence the use of fully automated electric vehicles to feed these higher capacity systems in a seamless way may be a good solution to bring more people to public transport and improve transport sustainability. The use of AVs for the last mile connection has been analysed before but mainly on a technology perspective (Chong et al., 2011). On the behaviour modelling side, Yap et al. (2016) positioned AVs as egress mode of train trips and explored the travellers' preferences for ATs. The authors applied a stated preference experiment to estimate a discrete choice model and

concluded that travellers' attitudes regarding vehicle automation are far from optimistic. However, as referred, the system they studied did not include AVs being used as conventional taxis in a city, meaning for all origin and destination pairs, which limits their findings for the purpose of informing this research.

In this chapter, we present an optimization approach to define the service area of an AT system which satisfies passengers' requests to access or egress a train station, in order to maximize the profit of the AT system. Since AVs can be relocated at a lower cost (no need to hire staff), the model considers the possibility of the vehicles travelling alone as a relocation method. Moreover, the system is based on mandatory pre-booking, allowing accepting or rejecting demand according to the profit maximization function. From a methodological point of view the models are based on the ones by Correia and Antunes (Correia and Antunes, 2012) , hence this study contributes to the literature by introducing a novel application of these formulations to the case where automated vehicles are used, thus avoiding the high costs that today the traditional carsharing operators have to consider.

A zoning problem is by definition a planning problem, however, to select trips is typically an operational problem which should be solved on a daily basis. In this chapter, we assume that our models are used on a daily basis for trip selection (operational purpose), but by running them with simulated trips for several replications before implementing the system we are able to obtain the zones which should be included in the service area around the station.

2.2 Integer programming models with two service schemes

In this section, we describe the formulation of two IP models in order to determine the optimal service area and trips to be served by an AT system. The two formulations depend on how trips are selected from the total number of reservations done in one typical day (24 hours in advance booking).

The first scheme (S1) is called free service. The model works on the assumption that the taxi company can achieve total control over trip selection, by being free to accept or reject requests according to the profit maximization. Waiting time is not applicable for the passengers since the trip is only served exactly at the starting time of the request. The model allocates each AT to a specific trip only if it will bring a higher daily profit. Otherwise, this request will be rejected with no extra penalty, even though there might be available taxis in the system. Such service scheme is flexible and profitable but will lead to unhappy customers because they may be in a situation in which they have their trips rejected but they know that some ATs are usually available nearby.

The second scheme (S2) is full service, which guarantees that all demand to/from selected zones must be satisfied. It does not mean that all the requests to/from potential zones will be met because zones are to be selected by the model too. Compared to S1, this scheme provides a favourable taxi service which assures that no requests will be missed from the served zones.

We consider in this chapter a first-mile/last-mile transport service for accessing and egressing from train stations, which means that we will not consider requests to use the system between two service zones where a service zone is a candidate area of the city for offering the transport service. Moreover, no requests will be considered when trips begin before the service period or when they end after that period. In order to provide a better service, we set up some relocation time before and after the service period and allow the taxis to move between the train station and service zones. At the beginning of the operation time, all vehicles are at the station where parking is free for the company. Similarly, taxis will all come back to the train station at the end of the operation period. This guarantees that the taxi fleet has enough time (normally during night time) to do the necessary full-charging of the battery and any required maintenance. When stopped in a service zone the car must pay for that parking.

Besides accepting or rejecting trips to and from the station, vehicles can travel empty from a service zone to another one to pick up another traveller. There is no need to have special equipment in a service zone because vehicles will only park there.

Before presenting the two models it is important to state all the assumptions which were considered in their formulations. To simplify we treat all origins and destinations of passengers' requests in the same zone as coming or going to the same point: the centroid of the service zone. Plus, we regard taxis as flows, which means that we do not differentiate a specific taxi. This hinders the computation of the specific battery charge that each vehicle has during the day, as modelled in (Correia and Santos, 2014), but it simplifies the problem to be solved.

Table 2.1 Notations

Notation	Description
Sets	
$\boldsymbol{N}$	$= \{0,1,\dots,i,\dots,N\}$, set of the train station plus the candidate service zones ($i = 0$ represents the train station).
$\boldsymbol{N'}$	$= \{1,\dots,i,\dots,N\}$, set of candidate service zones.
$\boldsymbol{T}$	$= \{0,1,\dots,t,\dots,T\}$, set of time instants in the service period. The time between two consecutive time instants is considered to be one time step where the number of time steps in a day is T.
$\boldsymbol{T'}$	$= \{-\delta_{max},\dots,-1,0,1,\dots,t,\dots,T,T+1,\dots,\delta_{max}\}$, set of time instants in the operation period, including the service period $\{0 \dots T\}$ and relocation period $\{-\delta_{max} \dots 0\}$ and $\{T \dots \delta_{max}\}$. δ_{max} is the maximum travel time between the train station and any potential zone.
Data & Parameters	
F	taxi fleet size in the system. The fleet size is an input of the model instead of a decision variable in order to guarantee the model linearity.
δ_{ij}	the travel time in time steps between zone i and zone j, $\forall i,j \in \boldsymbol{N}, i \neq j$.

δ_{max}	the maximum travel time in time steps between the train station and any service zone, $\delta_{max} = max\{\delta_{ij}\} \quad i = 0, \forall j \in \boldsymbol{N}' \; or \; i \in \boldsymbol{N}', j = 0$.
d_{ij}	the travel distance between zone i and zone j, $\forall i, j \in \boldsymbol{N}$, $i \neq j$.
$Q_{0i}^{t,t+\delta_{0i}}$	the number of passenger requests from the train station to zone i from time instant t to time instant $t + \delta_{0i}$, $\forall i \in \boldsymbol{N}'$, $\forall t \in \boldsymbol{T}$, $t + \delta_{0i} \leq \boldsymbol{T}$.
$Q_{i0}^{t,t+\delta_{i0}}$	the number of passenger requests from zone i to the train station from time instant t to time instant $t + \delta_{i0}$, $\forall i \in \boldsymbol{N}'$, $\forall t \in \boldsymbol{T}$, $t + \delta_{i0} \leq \boldsymbol{T}$.
R	battery range of a vehicle expressed in driving distance with a full battery (km).
E	distance that can be driven with a one time-step charging (km/time step).
P	price rate per driving distance (€ / km).
C_{m1}	vehicle maintenance costs per driving distance (€ / km).
C_d	depreciation cost per vehicle per day (€ / day).
C_p	parking price at the service zones per spot per time step (€ / spot × time step).
M	large number.
Decision variables	
x_i	equals to 1 if the candidate zone i can be served, otherwise 0, $\forall\, i \in \boldsymbol{N}'$.
$D_{0i}^{t,t+\delta_{0i}}$	the number of trips satisfied from the train station to service zone i from time instant t to time instant $t + \delta_{0i}$, $\forall i \in \boldsymbol{N}'$, $\forall t \in \boldsymbol{T}$, $t + \delta_{0i} \leq \boldsymbol{T}$.
$D_{i0}^{t,t+\delta_{i0}}$	the number of trips satisfied from service zone i to the train station from time instant t to time instant $t + \delta_{i0}$,$\forall i \in \boldsymbol{N}'$, $\forall t \in \boldsymbol{T}$, $t + \delta_{i0} \leq \boldsymbol{T}$.
$S_i^{t,t+1}$	the number of vehicles stocked at zone i from time instant t to time instant $t + 1$, $\forall i \in \boldsymbol{N}$, $\forall t \in \boldsymbol{T}'$, $t + 1 \leq T + \delta_{max}$.
$U_{ij}^{t,t+\delta_{ij}}$	the number of taxis travelling from zone i to zone j from time instant t to time instant $t + \delta_{ij}$,$\forall i, j \in \boldsymbol{N}$, $i \neq j$, $\forall t \in \boldsymbol{T}'$, $t + \delta_{ij} \leq T + \delta_{max}$.
V_i^t	the number of available vehicles at zone i at time instant t, $\forall i \in \boldsymbol{N}$, $\forall t \in \boldsymbol{T}'$.
Auxiliary variables	
Z_0	the number of parking spots at the train station.
σ_i	the total idle time in time steps that all taxis spend at service zone i in a day, $\forall i \in \boldsymbol{N}'$.
L_t	average driving distance per vehicle from the beginning of the day until time instant t, $\forall t \in \boldsymbol{T}$.
θ	share of satisfied demand (percentage).

In order to make this system feasible, a central management service is required to compute the best solution and give instructions to ATs. Thus all the requests have to be priorly known before the operation day via a reservation system. We argue that normally train trips are

planned in advance according to the timetable. This means that real-time booking is less likely to happen as much as in the Uber ("Uber," 2017) system. The notations used in this chapter are presented in Table 2.1.

2.2.1 Model S1- free service

Model S1 is the one with the free service scheme which allows the system to freely select trips according to the profit maximization. In this model, if there is at least one trip served in a zone, this zone is included in the total service area.

Considering the described system and its assumptions, we formulate the following integer programming problem:

Objective function:

$$Max\ \Pi = P \cdot \left(\sum_{\substack{i \in \boldsymbol{N'}, t \in \boldsymbol{T}, \\ t+\delta_{0i} \leq T}} D_{0i}^{t,t+\delta_{0i}} \cdot d_{0i} + \sum_{\substack{i \in \boldsymbol{N'}, t \in \boldsymbol{T}, \\ t+\delta_{i0} \leq T}} D_{i0}^{t,t+\delta_{i0}} \cdot d_{i0} \right) - C_{m1} \cdot \sum_{\substack{i,j \in \boldsymbol{N}, \\ i \neq j, t \in \boldsymbol{T'} \\ t+\delta_{ij} \leq T+\delta_{max}}} U_{ij}^{t,t+\delta_{ij}} \cdot d_{ij} - C_d \cdot F - C_{m2} \cdot Z_0 - C_p \cdot \sum_{i \in \boldsymbol{N'}} \sigma_i \tag{2.1}$$

The objective function (2.1) maximizes the total profit (Π) during a typical day of operations, taking into account the revenues paid by the passengers, vehicle maintenance costs, vehicle depreciation costs, parking space maintenance costs in the train station and parking costs in the service zones. There is no extra cost of rejecting requests in the system. At the same time, we use accepted and rejected requests percentage as an indicator to assess the performance of the AT system.

Constraints

$$V_0^{-\delta_{max}} = F \tag{2.2}$$

Constraint (2.2) describes the initial status of the AT fleet. It imposes that at the beginning of the operation period (before the beginning of the service period), all vehicles are stocked at the train station.

$$V_i^{-\delta_{max}} = 0 \quad \forall i \epsilon \boldsymbol{N'} \tag{2.3}$$

Constraints (2.3) impose that there cannot be any vehicle in the service zones at the beginning of the operation period.

$$V_0^{T+\delta_{max}} = F \tag{2.4}$$

Constraint (2.4) imposes that at the end of the operation period (after the end of the service period), all taxis come back to the train station.

$$V_i^{T+\delta_{max}} = 0 \quad \forall i \epsilon \boldsymbol{N}' \tag{2.5}$$

Constraints (2.5) guarantee that no taxi is available in any service zone at the end of the operation period.

$$S_i^{t,t+1} = S_i^{t-1,t} + \sum_{j \in \boldsymbol{N}, i \neq j} U_{ji}^{t-\delta_{ji},t} - \sum_{\substack{j \in \boldsymbol{N}, i \neq j, \\ t+\delta_{ij} \leq T+\delta_{max}}} U_{ij}^{t,t+\delta_{ij}} \quad \forall i \in \boldsymbol{N}, \forall t \in \boldsymbol{T}', t+1 \leq T + \delta_{max} \tag{2.6}$$

Constraints (2.6) yield the vehicle stock level at the train station or service zone i from time instant t to time instant $t+1$. The number of stocked taxis for time period from t to $t+1$ equals to the stocked taxis for time period from $t-1$ to t plus the vehicles coming into the train station or service zones minus the vehicles getting out.

$$V_i^{t+1} = V_i^t - \sum_{\substack{j \in \boldsymbol{N}, i \neq j, \\ t+\delta_{ij} \leq T+\delta_{max}}} U_{ij}^{t,t+\delta_{ij}} + \sum_{j \in \boldsymbol{N}, i \neq j} U_{ji}^{t+1-\delta_{ji},t+1} \quad \forall i \in \boldsymbol{N}, \forall t \in \boldsymbol{T}', t+1 \leq T + \delta_{max} \tag{2.7}$$

Constraints (2.7) are flow conservation constraints which compute the number of available taxis at node i_{t+1} (the train station or service zone i at time instant $t+1$) as a function of the number of taxis at time instant t minus the vehicles getting out plus the vehicles coming into the train station or service zones at next time instant.

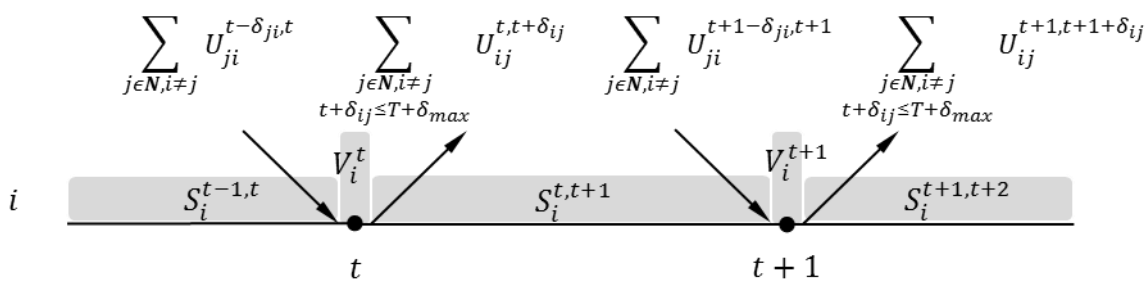


Figure 2.1 Vehicle movement in zone *i*

We use $S_i^{t,t+1}$ to describe those taxis who are waiting from one time instant to the next time instant (a whole time step) at the train station or a service zone. During this time, taxis are waiting for the next trip or charging to get electric power (if they are at the train station). V_i^t is used to represent the instantaneous number of vehicles at a time instant t. We assume that the vehicles become available at a time instant t after they arrive at that point and before they go out to the next destination. This can be seen in **Figure 2.1**.

$$\sigma_i = \sum_{\substack{t \in \boldsymbol{T}' \\ t+1 \leq T+\delta_{max}}} S_i^{t,t+1} \qquad \forall i \in \boldsymbol{N}' \tag{2.8}$$

Constraints (2.8) compute the total parking time in time steps for the whole day in each service zone. To be more specific, $S_i^{t,t+1}$ is the stock of vehicles, so it corresponds to the number of parking spots occupied for one time step in each service zone i in that period. Summing them up yields the total parking time which will be used to calculate the parking costs in each service zone.

$$V_0^t \leq Z_0 \qquad \forall t \in \boldsymbol{T}' \tag{2.9}$$

Constraints (2.9) ensure that the parking capacity in the train station is enough for the number of ATs present there at any time instant.

$$D_{0i}^{t,t+\delta_{0i}} \leq Q_{0i}^{t,t+\delta_{0i}} \cdot x_i \qquad \forall i \in \boldsymbol{N}', \forall t \in \boldsymbol{T}, t+\delta_{0i} \leq T \tag{2.10}$$

Constraints (2.10) assure that the satisfied trips between the train station and service zone i which begin at time instant t and finish at time instant $t+\delta_{0i}$ tmust be lower than or equal to the passengers' requests on the same OD. And if zone i cannot be served (x_i=*0*), the satisfied demand must be zero.

$$D_{i0}^{t,t+\delta_{i0}} \leq Q_{i0}^{t,t+\delta_{i0}} \cdot x_i \qquad \forall i \in \boldsymbol{N}', \forall t \in \boldsymbol{T}, t+\delta_{i0} \leq T \tag{2.11}$$

Constraints (2.11) assure that the satisfied trips between service zone i and the train station which begin at time instant t and finish at time instant $t+\delta_{i0}$ must be lower than or equal to the passengers' requests on the same OD. And if zone i cannot be served (x_i=0), the satisfied demand must be zero.

$$\sum_{\substack{j \in \boldsymbol{N}, t \in \boldsymbol{T}', \\ i \neq j, \\ t+\delta_{ij} \leq T+\delta_{max}}} U_{ij}^{t,t+\delta_{ij}} + \sum_{\substack{j \in \boldsymbol{N}, t \in \boldsymbol{T}', \\ i \neq j, \\ t+\delta_{ji} \leq T+\delta_{max}}} U_{ji}^{t,t+\delta_{ji}} \leq M \cdot x_i \qquad \forall i \in \boldsymbol{N}' \tag{2.12}$$

Constraints (2.12) guarantee that if a zone cannot be served, there are no taxis travelling between this zone and any other zones or the train station.

$$x_i \leq \sum_{\substack{t \in \boldsymbol{T}, \\ t+\delta_{0i} \leq T}} D_{0i}^{t,t+\delta_{0i}} + \sum_{\substack{t \in \boldsymbol{T}, \\ t+\delta_{i0} \leq T}} D_{i0}^{t,t+\delta_{i0}} \quad \forall i \in \boldsymbol{N}' \tag{2.13}$$

Constraints (2.13) assure that if no trip from zone i is satisfied by the ATs then that zone is not selected.

$$D_{0i}^{t,t+\delta_{0i}} \leq U_{0i}^{t,t+\delta_{0i}} \qquad \forall i \in \boldsymbol{N}', \forall t \in \boldsymbol{T}, t+\delta_{0i} \leq T \tag{2.14}$$

Constraints (2.14) impose the condition that the number of vehicles travelling between the train station and service zone i must be greater than or equal to the number of people

travelling on that OD (note that the vehicle capacity is just one seat) since the vehicles can travel without any human inside.

$$D_{i0}^{t,t+\delta_{i0}} \leq U_{i0}^{t,t+\delta_{i0}} \quad \forall i \in \boldsymbol{N}', \forall t \in \boldsymbol{T}, t+\delta_{i0} \leq T \tag{2.15}$$

Constraints (2.15) impose the condition that the number of vehicles travelling between service zone i and the train station must be greater than or equal to the number of people travelling on that OD since the vehicles can travel without any human inside.

$$\sum_{\substack{j \in \boldsymbol{N}, i \neq j, \\ t+\delta_{ij} \leq T+\delta_{max}}} U_{ij}^{t,t+\delta_{ij}} \leq V_i^t \quad \forall i \in \boldsymbol{N}, \forall t \in \boldsymbol{T}' \tag{2.16}$$

Constraints (2.16) guarantee that the vehicles leaving a service zone at time instant t is less than or equal to the available vehicles at that service zone i.

$$\theta = \sum_{i \in \boldsymbol{N}'} \Big(\sum_{\substack{t \in \boldsymbol{T}, \\ t+\delta_{0i} \leq T}} D_{0i}^{t,t+\delta_{0i}} + \sum_{\substack{t \in \boldsymbol{T}, \\ t+\delta_{i0} \leq T}} D_{i0}^{t,t+\delta_{i0}} \Big) / \sum_{i \in \boldsymbol{N}'} \Big(\sum_{\substack{t \in \boldsymbol{T} \\ t+\delta_{0i} \leq T}} Q_{0i}^{t,t+\delta_{0i}} + \sum_{\substack{t \in \boldsymbol{T} \\ t+\delta_{i0} \leq T}} Q_{i0}^{t,t+\delta_{i0}} \Big) \tag{2.17}$$

Constraint (2.17) yields the percentage of satisfied demand.

$$L_t = \sum_{\substack{i \in \boldsymbol{N}, j \in \boldsymbol{N}, \\ i \neq j, t_1 \in \boldsymbol{T}', \\ t_1+\delta_{ij} \leq t}} U_{ij}^{t_1,t_1+\delta_{ij}} \cdot d_{ij} / F \quad \forall t \in \boldsymbol{T}' \tag{2.18}$$

Constraints (2.18) compute the average driving distance per vehicle from the beginning of the operation period to a particular time instant t.

$$L_t - R \leq \sum_{\substack{t_1 \in \boldsymbol{T}', \\ t_1+1 \leq t}} S_0^{t_1,t_1+1} \cdot E / F \quad \forall t \in \boldsymbol{T}' \tag{2.19}$$

Constraints (2.19) impose the charging condition i.e. when the average driving distance per vehicle exceeds the driving range of a full battery, taxis must stay long enough at the train station to get sufficient power for the remaining average driving distance.

$$V_i^t \geq 0 \quad \forall i \in \boldsymbol{N}, \forall t \in \boldsymbol{T}' \tag{2.20}$$

$$S_i^{t,t+1} \geq 0 \quad \forall i \in \boldsymbol{N}, \forall t \in \boldsymbol{T}', t+1 \leq T+\delta_{max} \tag{2.21}$$

$$D_{0i}^{t,t+\delta_{0i}} \geq 0 \quad \forall i \in \boldsymbol{N}', \forall t \in \boldsymbol{T}, t+\delta_{0i} \leq T \tag{2.22}$$

$$D_{i0}^{t,t+\delta_{i0}} \geq 0 \quad \forall i \in \boldsymbol{N}', \forall t \in \boldsymbol{T}, t+\delta_{i0} \leq T \tag{2.23}$$

$$U_{ij}^{t,t+\delta_{ij}} \geq 0 \quad \forall i,j \in \boldsymbol{N}, i \neq j, \forall t \in \boldsymbol{T}', t + \delta_{ij} \leq T + \delta_{max} \tag{2.24}$$

$$Z_0 \geq 0 \tag{2.25}$$

$$x_i = (0,1) \quad \forall i \in \boldsymbol{N}' \tag{2.26}$$

Constraints (2.20) - (2.26) define the domain for the decision variables.

2.2.2 Model S2- full service

This model considers that all requested trips to/from served zones must be satisfied. The model is based on the previous by adding the following constraints:

$$D_{0i}^{t,t+\delta_{0i}} \geq Q_{0i}^{t,t+\delta_{0i}} + M \cdot (x_i - 1) \quad \forall i \in \boldsymbol{N}', \forall t \in \boldsymbol{T}, t + \delta_{0i} \leq T \tag{2.27}$$

$$D_{i0}^{t,t+\delta_{i0}} \geq Q_{i0}^{t,t+\delta_{i0}} + M \cdot (x_i - 1) \quad \forall i \in \boldsymbol{N}', \forall t \in \boldsymbol{T}, t + \delta_{i0} \leq T \tag{2.28}$$

Constraints (2.27) and (2.28) guarantee that when zone i is selected as service zone (x_i=1), the satisfied trips will be equal to or more than the requests. This means all requests generated to/from zone i are satisfied because working with constraints (2.10) and (2.11) will impose $D_{0i}^{t,t+\delta_{0i}} = Q_{0i}^{t,t+\delta_{0i}}$ and $D_{i0}^{t,t+\delta_{i0}} = Q_{i0}^{t,t+\delta_{i0}}$.

2.2.3 Bounding the problem

The search for the optimum set of trips to be satisfied can be accelerated by bounding the problem with extra constraints. They do not change the solution region but tight the bounds on that space by eliminating non-integer solutions of the relaxation process of the traditional branch and bound search method. Thus the gap to find the optimal solution closes faster since certain nodes of the tree will not be explored.

$$S_i^{t,t+1} \leq F \cdot x_i \quad \forall i \in \boldsymbol{N}', \forall t \in \boldsymbol{T}', t + 1 \leq T + \delta_{max} \tag{2.29}$$

$$V_i^t \leq F \cdot x_i \quad \forall i \in \boldsymbol{N}', \forall t \in \boldsymbol{T}' \tag{2.30}$$

These constraints impose that zones can only have taxis there when they are selected as being part of the service area. These constraints are already imposed by the interaction of constraints (2.6), (2.7) and (2.12). In the branch and bound process, it may happen that the relaxed variables on the left-hand side of constraints (2.6) and (2.7) are equivalent to the right-hand side, but in reality, not all of them may have a value greater than 0 at the same time. Imposing bound constraints (2.29) and (2.30) will not allow those solution nodes to be explored in the relaxation process.

2.2.4 A small scale example

We apply model S1 and model S2 to a small scale example to illustrate how they work and the type of results which can be obtained. In this example, we consider passenger requests

between 4 potential zones and the train station. There are eight time steps from time instant 0 to 8 in the service period. As the maximum travel time between any zone and the train station is three time steps, we set the operation period from time instant -3 to 11. Table 2.2 and Figure 2.2 show the passenger requests in the service period. Each line corresponds to one request. The trip duration is based on the travel time between two points which is a set of parameters in this model. For example, we can see that there is a request from the train station to zone 4 starting at time instant 0 and the trip duration is three time steps.

Table 2.2 Passenger requests for the small scale example

Starting time	Origin	Destination	Number of requests
0	0	4	1
0	0	2	1
2	0	3	1
3	0	2	1
4	0	1	1
6	0	3	1
6	0	2	1
0	1	0	1
2	3	0	1
2	4	0	1
3	2	0	1
4	4	0	1
6	1	0	1

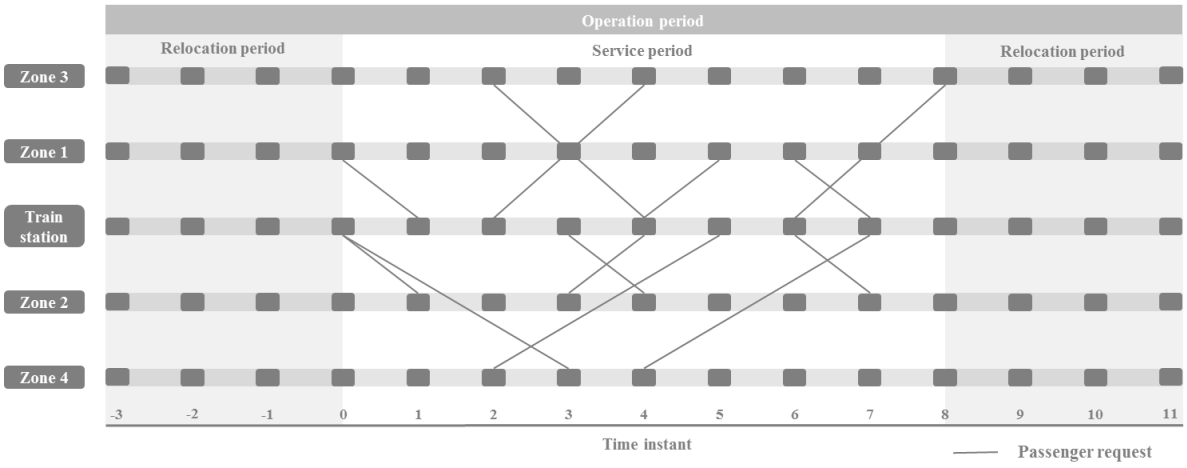


Figure 2.2 Passenger requests for the small scale example

Table 2.3 presents the travel time and travel distance between the train station and potential zones. For this small example, we consider the following parameters: price rate P=1 euro/km, vehicle maintenance cost C_{m1} =0.05 euro/km, vehicle depreciation cost C_d =17 euro/day parking space maintenance C_{m2}=5 euro/spot/day and parking price C_p=0.25 euro/spot/time step.

Table 2.3 Travel time and travel distance

	Travel time (time step)					Travel distance (km)				
Zones	0	1	2	3	4	0	1	2	3	4
0	-	1	1	2	3	-	2	3	4	6
1	1	-	2	1	3	2	-	4	2	6
2	1	2	-	1	2	3	4	-	2	3
3	2	1	1	-	2	4	2	2	-	4
4	3	3	2	2	-	6	6	3	4	-

In Figure 2.3 we show the graphical results of model S1 for vehicle movements ($U_{ij}^{t,t+\delta_{ij}}$) in the optimal solution and the number of available ATs stocked at each zone and at the train station (V_i^t). Beyond the service period, the ATs are allowed to travel between the train station and the service zones. In this case, two taxis do relocations in the first relocation period and another one in the second relocation period. Without this, the system cannot serve trips like $D_{10}^{0,1}$ or getting back to the train station for daily maintenance. Together with the demand in Figure 2.2, each solid line represents the accepted request ($D_{0i}^{t,t+\delta_{0i}}$ or $D_{i0}^{t,t+\delta_{i0}}$). The dashed lines are the taxi relocations ($U_{ij}^{t,t+\delta_{ij}} - D_{0j}^{t,t+\delta_{0j}}$ or $U_{ij}^{t,t+\delta_{ij}} - D_{i0}^{t,t+\delta_{i0}}$) which are vehicle movements without passengers. Moreover, a line from Figure 2.2 which disappears in Figure 2.3, indicates that a request from zone 2 to the train station starting at time instant 3 is rejected in the optimal solution.

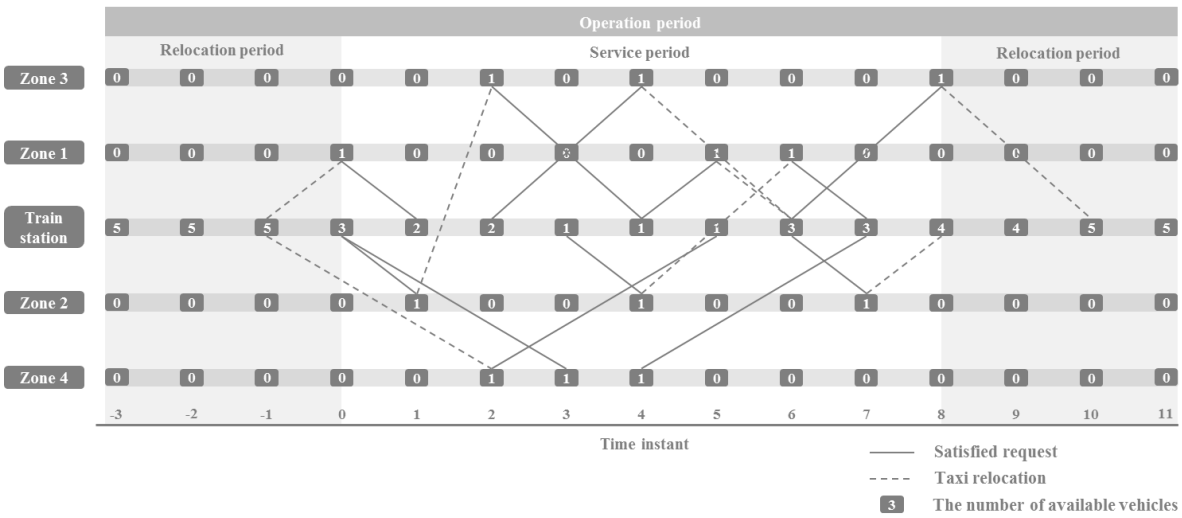


Figure 2.3 Results from model S1 for the small scale example

Model S2 is also applied to the small scale example, with optimal results shown in Figure 2.4. With the service rule imposing that all requests from selected zones must be satisfied, the model decided to serve zone 1, zone 2 and zone 4 (all requests to/from these zones being satisfied). However, no trip in zone 3 means this zone is apparently not leading to maximizing the profit.

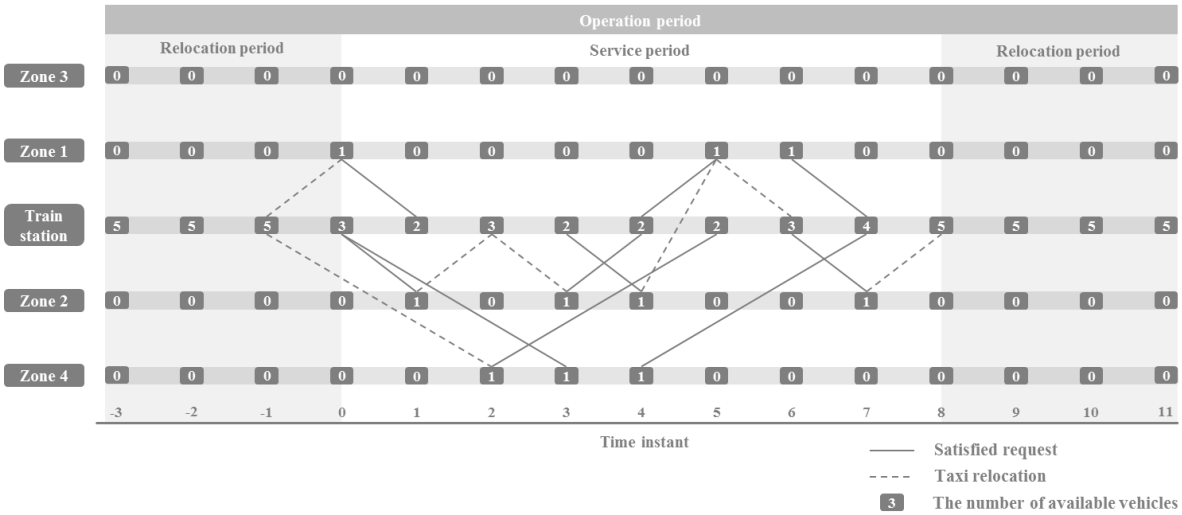


Figure 2.4 Results from model S2 for the small scale example

2.3 Case study

We apply our model to a bigger case study city, the city of Delft, which is located in the Netherlands. The municipality has a total area of 24 km2 and a population of about 100,000. The city has two train stations but we focus on the one with less demand, Delft Zuid, which is located on a section of the railway line between The Hague and Rotterdam. It counts 142 sprinter trains per day, which provide slower stopping train service not only at significant stations but also at smaller stations between cities (NS, 2014). This station is connected to the city area of Delft by just one regional bus line hence its main purpose is local accessibility where the TU Delft campus is one of the most significant trip attractors/generators.

The demand is mainly concentrated in two peak periods: 7:00-9:00 am and 4:00-6:00 pm. The average headway in the two directions is 15 minutes. Moreover, given that the first train leaves at 5:27 am and the last train at 1:01 am in the next day, we choose 6:00 to 22:00 as the service period of the AT service to provide accessibility to the train passengers. The time step will be 5 minutes in order to reduce computational time. Thus for the 16 hours service period in a day, there will be 192 time steps considered from time instant 0 to time instant 192. The maximum travel time between the train station and the service zone is 3 time steps hence the model operation period is defined from time instant -3 to time instant 195.

We divided the catchment area of the Delft Zuid station into 48 potential service zones trying to follow a principle of homogeneous land use in each zone (Figure 2.5). The average size of each zone is 500m×500m.

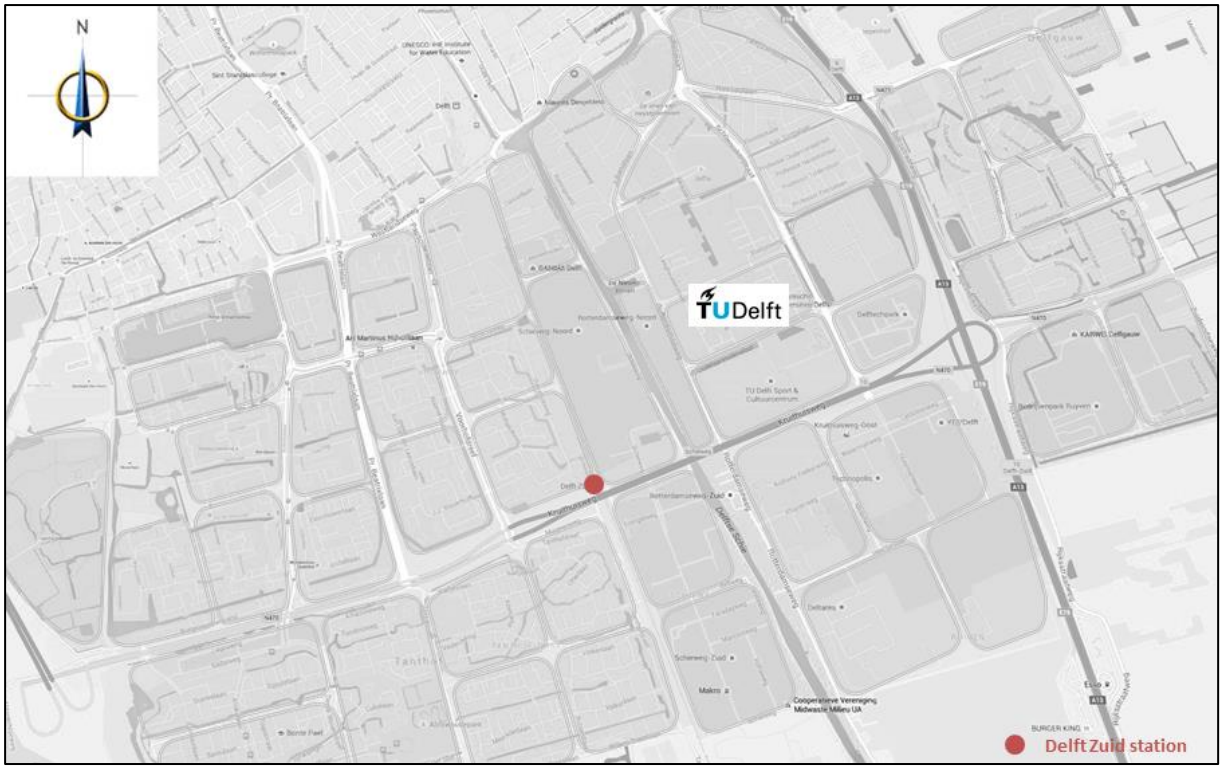


Figure 2.5 Potential service zone location.

For the purpose of applying model S1 and S2 for studying the AT system in Delft Zuid station, we need these data: 1) all requests during an average day; 2) costs of running the system; 3) driving distance and travel time between the train station and any potential service zone.

In order to determine the demand, we conducted a face to face field survey at the train station on the 2nd and 4th of June, 2015. This survey allowed obtaining a stated preference of the passengers for using the ATs, and also an estimation of the probability of each potential zone being an origin or a destination in these rail connection trips. Regarding the starting time of each request, we estimate that the passenger arriving time at the train station coming from the city is uniformly distributed. Then in order to get the departure time from the service zone, we subtract the determinist taxi travel time in that OD pair. Correspondingly, we assume that the passengers who just finished the railway trip will leave the train station to go to the city area immediately. Consequently, there is an equal probability of passengers to leave the station at the end of each train headway during a day.

The total number of passengers using Delft Zuid station in a working day is 4668, a number provided by NS (the passenger railway operator in the Netherlands) for an average day in 2013. From the data we see that 50% of the passenger are arriving at the station and the other 50% leaving the station, yielding perfect balance. Each trip was allocated to either using an AT or another transport mode based on the survey results.

In order to generate the AT reservations, we use the Monte Carlo simulation, where we select an origin, a destination, and a starting time for each trip based on the estimated probabilities brought from the survey. Using a random seed, we get a new set of AT passenger requests each time when a replication is generated.

Travel distance d_{ij} and travel time δ_{ij} between any potential service zones and the train station were calculated using the web mapping service Google Maps. These two inputs are computed between each centroid of a zone, where travel times will be considered deterministic which means they will not change throughout the day as mentioned before.

The cost values considered are as follows:

- P (price rate): 1 euro/km.
- C_{m1} (vehicle maintenance cost) : 0.05 euro/km.
- C_d (depreciation cost): 17 euro/day.
- C_{m2} (parking space maintenance): 5 euro/spot/day.
- C_p (parking price in service zones): 0.25 euro/spot/time step.

We use as reference the price rate of Uber in Amsterdam and Rotterdam in the Netherlands ("Uber," 2017). The fare is 1.4 € / km together with 2 € as base fare. ATs do not need drivers which makes the operating cost lower. Besides, a lower fare will attract more passengers to use this brand new transport mode.

We chose Renault Twizy for our one seat capacity vehicle. The battery range of this electric vehicle is 80 km and normally the charging time is 4 hours ("Discover Renault Twizy," 2014). Based on this and the assumption that the battery reached by charging is uniform, the charging efficiency is 1.67 km/time step in this case. Therefore, the charging parameters considered are as follows:

- R (battery range): 80 km.
- E (the distance that can be driven with a one time-step charging): 1.67 km/time step.

2.4 Experiments and results

Several experiments were done with the S1 and S2 trip selection schemes. The optimization models were run in an i7 processor @2.10GHz, 8.00GB RAM computer under a Windows 7 64-bit operation system by Xpress, an optimization tool which uses advanced branch-and-cut algorithms for solving integer programming problems (FICO, 2014).

The passenger request generation is done, as referred, through Monte Carlo simulation. The detailed information of each trip including if it is going to be taken by an AT, its origin, its destination and the starting time are randomly generated thus we have decided to run 10 replications namely 10 sets of passenger requests. The final results of each model run are presented as an average value for each indicator.

The average total number of reservation requests for the whole day with the10 replications was 2061. These were associated with 44 potential zones, which indicates that 4 zones do not have any requests.

2.4.1 Fleet Size Variation

Four scenarios were built with a fleet of 20, 40, 60 and 80 ATs for model S1 and S2. The computation time of model S1 is less than 3 minutes for one replication and a bigger fleet size leads to faster running time. For model S2 the running time of 20 vehicles is about 16 minutes and for 40 vehicles it is 9 minutes. But for 60 and 80 vehicles, one replication run is less than 1 minute. The results obtained for scheme S1 are presented in Table 2.4.

Table 2.4 Optimization results for model S1 for fleet size variation

Total requests	Fleet size	Profit (euro/day)	Total service zones	Requests satisfied	Requests satisfied (%)	Driving distance /Veh.* (km)	Driving time /Veh. (h)	Satisfied trips /Veh.
2061	20	2365.3	43	1242	60.2%	204.8	9.4	62
	40	3391.3	44	1955	94.9%	185.0	8.9	49
	60	3128.0	44	2059	99.9%	129.4	6.3	34
	80	2691.2	44	2061	100.0%	97.1	4.7	26

Veh.*: vehicle

For model S1 which is the free level of service, in a scenario with 20 taxis, only 60.2% of the potential requests are satisfied (1242 out of 2061) generating a maximum daily profit of 2365.3 €. The three indicators of taxi movement show that vehicles in this scenario have a higher usage: each taxi drives an average of 204.8 km, 9.4 hours, serving 62 trips in a day. From the scenario with 40 taxis in the system, we can see that the satisfied requests increases to 94.9%. Having more ATs not only increases the number of served trips but also improves the daily profit to 3391.3 €, which is the highest value among the four scenarios in model S1. When the number of ATs goes to 60, the percentage of accepted requests reaches 99.9%. Only 2 passenger requests are rejected. In this scenario, the taxi usage is much lower than the first one, with 129.4 km average driving and 34 trips served by each vehicle. With regards to the fleet of 80 taxis, it is easy to find that the optimization of service zone location and trip selection covers all the demand and all the zones having passenger requests. Although there are more requests satisfied by the system in this scenario, which denotes a higher level of service, the daily profit is lower than the best scenario, even lower than the third scenario because of the increase of depreciation cost associated to larger fleet size.

The number of served zones for the first scenario is 43 and for the remaining is always 44. Hence there always 43 zones being chosen as service zones regardless of the number of trips served from each zone. When the ATs are insufficient for the requests (fleet size 20), the model indeed closes one zone because the system should not serve any request from it.

Using the same demand, we applied the trip selection scheme S2. Results are presented in Table 2.5.

Table 2.5 Optimization results for model S2 for fleet size variation

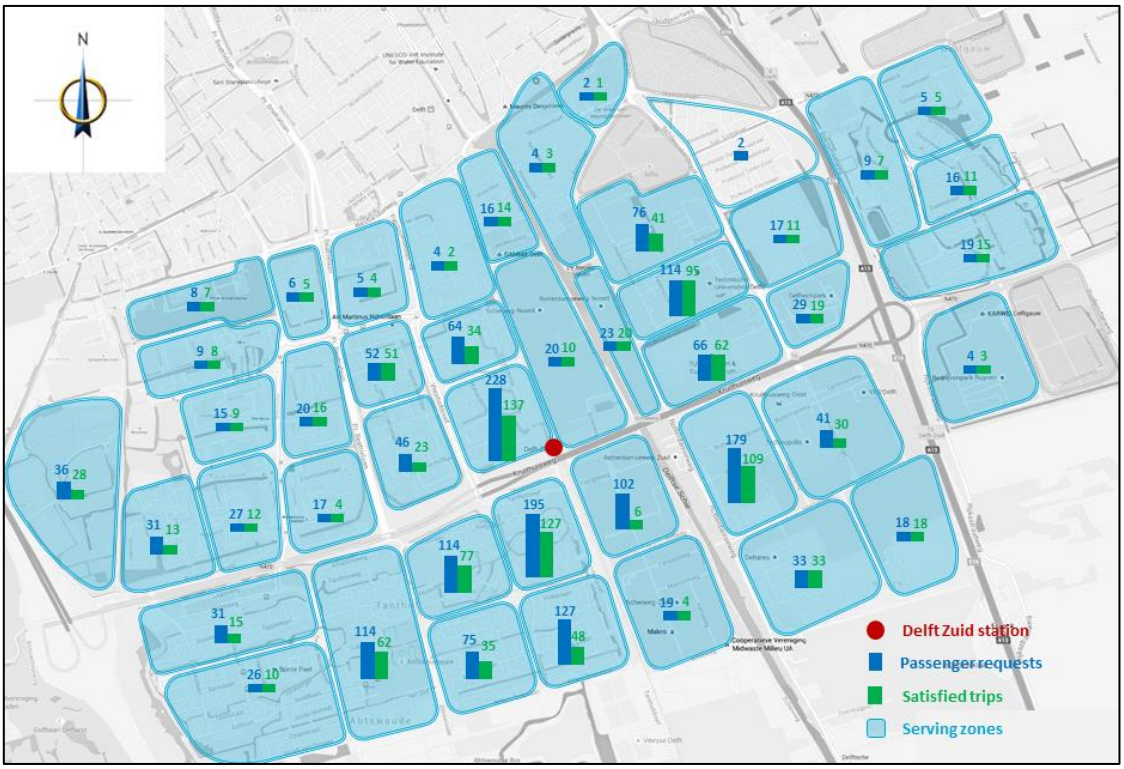
Total requests	Fleet size	Profit (euro/day)	Total service zones	Requests satisfied	Requests satisfied (%)	Driving distance /Veh.* (km)	Driving time /Veh. (h)	Satisfied trips /Veh.
2061	20	1645.8	21	913	44.3%	188.2	8.3	46
	40	2716.3	35	1648	80.0%	157.6	7.4	41
	60	3078.2	43	2034	98.7%	128.0	6.2	34
	80	2691.2	44	2061	100.0%	97.1	4.7	26

Veh.*: vehicle

For the scenario with a fleet of 20 ATs, only 21 zones and 913 requests are served which is significantly less than the same scenario in model S1. When increasing the fleet of vehicles in the system, the service area expands and the satisfied trips increase. It is improved by using 40 vehicles as ATs resulting in 35 zones on average for which all trips are satisfied. But this is still less than the same scenario in model S1. Similarly, this change can also be seen in the third scenario. When 80 taxis are available, the satisfied demand reaches 100%, with all potential zones which have requests being served. In this model, the most profitable scenario is the one with 60 ATs. A possible reason is that 40 ATs can serve more trips but they cannot capture all the requests from these zones, which makes these vehicles redundant to the system and leads to less profit for the taxi company.

The results for the service area optimization are shown in Figure 2.6 and Figure 2.7. For trip selection scheme S1, the model selects the potential trips to serve, therefore, the zones with satisfied trips, even if just one, are selected to comprise the service area. There are 44 zones with requests out of the 48 potential ones. When increasing the fleet size from 20 to 40, more trips from each zone are served by AVs and some zones are actually fully served.

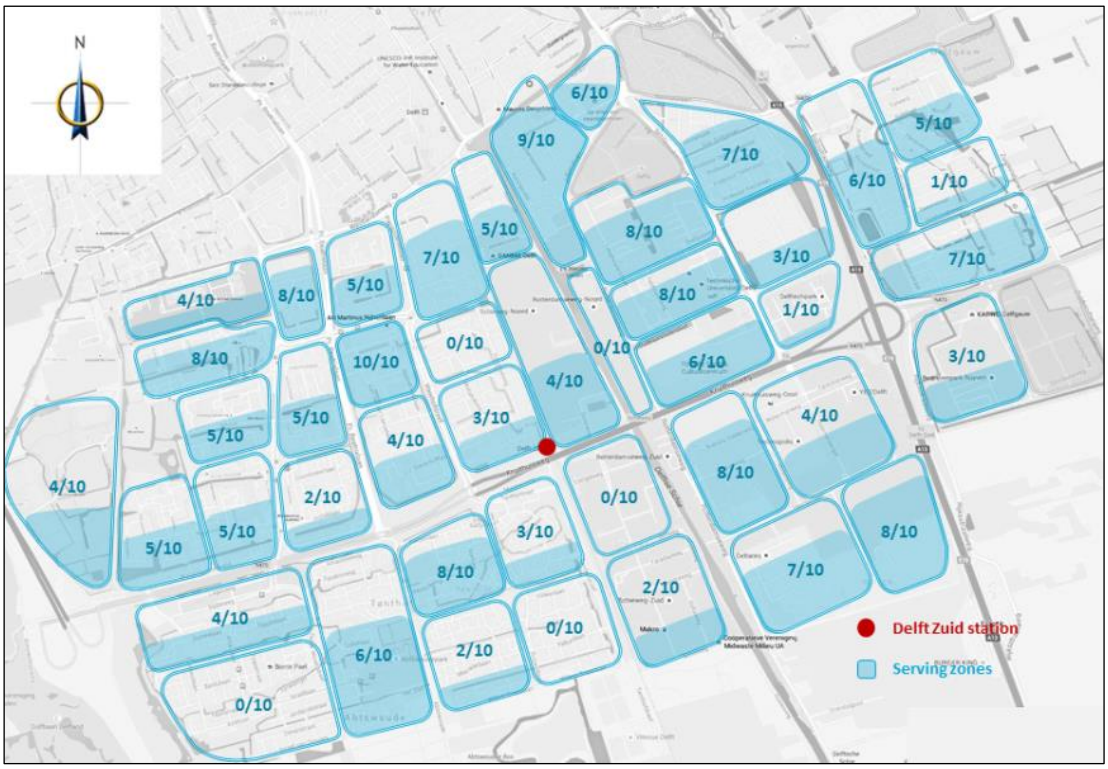
For model S2, the filling colour level in each zone indicates the number of replications when this zone was selected. In the scenario with 20 AVs, there are 4 zones which are never selected to be part of the service area. Results for the other areas are very diverse. For example, the zone on the west side is served in 4 replications and in the other 6 replications, the model decided to reject all trips to/from that zone. This shows that there is an important influence of the trip stochasticity. In the scenario with 40 vehicles, all zones have more replications in which they are selected. Moreover, there are 7 zones which are chosen in every replication, which indicates that these should always be part of the system.



(a) Scheme S1, 20 AVs Scenario

(b) Scheme S1, 40 AVs Scenario

Figure 2.6 Optimization results for scheme S1



(a) Scheme S2, 20 AVs Scenario

(b) Scheme S2, 40 AVs Scenario

Figure 2.7 Optimization results for scheme S2

In order to verify if 10 replications are sufficient to remove the noise of the simulation results, we compute the coefficient of variation (CV) of each indicator for 10 runs. Results are shown in Table 2.6.

Table 2.6 Coefficient of variation of each indicator for Model S1 and S2

Fleet size	S1			S2		
	Profit (€/day)	Requests satisfied (Trips)	Total service zones	Profit (€/day)	Requests satisfied (Trips)	Total service zones
20	0.97%	1.08%	2.89%	3.93%	4.22%	10.92%
40	1.79%	1.41%	1.71%	3.76%	3.76%	6.14%
60	2.25%	1.56%	1.71%	2.42%	1.97%	3.53%
80	2.82%	1.74%	1.71%	2.82%	1.74%	1.71%

We can see that almost all indicators have a relatively low level of CV, which means that under the requests variability over 10 simulations the model results are stable. The bigger value of CV is the one with 20 ATs in Model S2. This may be because the ATs are insufficient to serve all the requests.

2.4.2 Electric vs Conventional Vehicles

In the models, constraints (2.18) and (2.19) are used to compute the average driving distance and guarantee that on average the charging time is enough to power all the vehicles. Here we do a comparison between conventional ATs and electric ATs to test the impact of different technologies on the profit of the system. Though we assume that in the future the technology of electric vehicles will be mature which makes the vehicle maintenance costs the same as in conventional vehicles. Therefore in this computing the parameter C_{m1} (vehicle maintenance cost) remains the same.

As the scenario with 20 taxis has the most efficient vehicle usage and the scenario with 40 ATs has the best daily profit, we choose 20 and 40 as fleet size to run the model without constraints (2.18) and (2.19) for both trip selection schemes.

Table 2.7 Optimization results for electric ATs and conventional ATs

Model	Vehicle type	Fleet size	Profit (€/day)	Requests satisfied (%)	Model	Vehicle type	Fleet size	Profit (€/day)	Requests satisfied (%)
S1	Ele.*	20	2365.3	60.20%	S2	Ele.	20	1645.8	44.30%
	Con.**		3071.0	76.60%		Con.		1645.8	44.30%
	Ele.	40	3391.3	94.90%		Ele.	40	2716.3	80.00%
	Con.		3391.3	94.90%		Con.		2716.3	80.00%

Ele.*: electric
Con.**: Conventional

Table 2.7 shows the optimal results of using electric and conventional taxis. In model S1, when there are 20 taxis available in the system, the conventional ATs can serve more trips than the electric ATs. The different results prove that constraints (2.18) and (2.19) constrain the feasible region and influence the optimal solution. This also shows the fact that conventional ATs can serve trips continuously and electric vehicles cannot play the same role on the system performance. Regarding the 40 taxis fleet, there is no difference between the two optimization results thus, in this case, the stocking time at the train station is long enough on average for the vehicles to get enough power for the trips.

Model S2 has a different outcome for this comparison. It is easy to see that the optimal results are the same between the scenario with electric taxis and with conventional taxis when the fleet size is 20 and 40. This shows that constraints (2.27) and (2.28) have more significant impacts on the feasible region than constraints (2.18) and (2.19) therefore taking out the constraints about electric vehicle does not influence the optimal solution of this model. From a practical point of view, these results show that the ATs have enough time to get charged when they have to serve all the requests in a service area.

2.4.3 Automated taxis vs human-driven taxis

Automated taxis can drive themselves between any two zones when they are empty. But if the taxi is not automated, the company needs to hire staff to drive conventional taxis for vehicle relocation like in any carsharing system (Correia and Antunes, 2012; Jorge et al., 2014). In this section, we do a comparison between automated taxis and human-driven taxis on the daily profit thus exploring the benefits of using ATs.

For the objective function (2.1), staff costs should now be added. The new objective function is hence as follows:

$$
\begin{aligned}
Max\ \Pi = P \cdot \left(\sum_{\substack{i \in \boldsymbol{N'}, t \in \boldsymbol{T}, \\ t+\delta_{0i} \le T}} D_{0i}^{t,t+\delta_{0i}} \cdot d_{0i} + \sum_{\substack{i \in \boldsymbol{N'}, t \in \boldsymbol{T}, \\ t+\delta_{i0} \le T}} D_{i0}^{t,t+\delta_{i0}} \cdot d_{i0} \right) - C_{m1} \\
\cdot \sum_{\substack{i,j \in \boldsymbol{N}, \\ i \ne j, t \in \boldsymbol{T'} \\ t+\delta_{ij} \le T+\delta_{max}}} U_{ij}^{t,t+\delta_{ij}} \cdot d_{ij} - C_d \cdot F - C_{m2} \cdot Z_0 - C_p \cdot \sum_{i \in \boldsymbol{N'}} \sigma_i - C_h \\
\cdot \Big(\sum_{\substack{i,j \in \boldsymbol{N}, \\ i \ne j, t \in \boldsymbol{T'} \\ t+\delta_{ij} \le T+\delta_{max}}} U_{ij}^{t,t+\delta_{ij}} \cdot \delta_{ij} - \sum_{\substack{i \in \boldsymbol{N'}, t \in \boldsymbol{T}, \\ t+\delta_{0i} \le T}} D_{0i}^{t,t+\delta_{0i}} \cdot \delta_{0i} \\
- \sum_{\substack{i \in \boldsymbol{N'}, t \in \boldsymbol{T}, \\ t+\delta_{i0} \le T}} D_{i0}^{t,t+\delta_{i0}} \cdot \delta_{i0} \Big)
\end{aligned}
\tag{2.31}
$$

where C_h is the cost of a driver per time step (€ / time step).

According to the regulations in the Netherlands, the minimum wage for employees is 9 euro/hour, thus in this model we set C_h=0.75 euro/ time step.

Table 2.8 shows the results from applying the optimization model using the automated driving mode and the human-driven mode when relocating taxis between zones and train station.

Table 2.8 Optimization results for automated relocation and human-driven relocation

S1					S2				
Fleet size	Relocation	Profit (€/day)	Requests satisfied	Requests satisfied (%)	Fleet size	Relocation	Profit (€/day)	Requests satisfied	Requests satisfied (%)
20	Aut*	2365.3	1242	60.2%	20	Aut	1645.8	913	44.3%
	HuD**	1948.4	1242	60.2%		HuD	1150.8	908	44.1%
40	Aut	3391.3	1955	94.9%	40	Aut	2716.3	1648	80.0%
	HuD	2383.7	1943	94.3%		HuD	1864.5	1670	81.1%
60	Aut	3128.0	2059	99.9%	60	Aut	3078.2	2034	98.7%
	HuD	2036.4	2042	99.1%		HuD	1994.4	2034	98.7%
80	Aut	2691.2	2061	100.0%	80	Aut	2691.2	2061	100.0%
	HuD	1598.1	2044	99.2%		HuD	1594.8	2061	100.0%

Aut*: Automated
HuD**: human-driven

In model S1, there is no difference between automated driving and human-driven relocations regarding the number of served trips when there are 20 taxis in the system. But the daily profit of manually driven cars is lower than in the automation mode which reflects the most significant benefits of using automated taxis. In the scenarios with 40, 60 and 80 taxis, the optimal solutions change in relation to the number of satisfied trips. Besides, the cost for the staff, together with less satisfied trips, makes the profit of the taxi company lower than when the vehicles are automatic. This happens because the system decides to reject some trips to avoid more driving costs. Even though accepting these trips will bring revenues, it cannot cover the high relocation cost.

In model S2, the optimal solutions between automated or human-driven taxis only change in the scenario with 20 and 40 taxis. In the 60 and 80 vehicles' fleet, the satisfied trips remain the same number and the profit is lower due to the driver costs. This indicates that when taxis are sufficient for this case study (scenarios with 60 and 80), constraints (2.27) and (2.28) in mandatory trip satisfaction have a stronger influence on the feasible region of solutions than adding driver costs.

2.5 Conclusions

AVs have gained great momentum in the recent decades evolving from concept vehicles to starting to drive in our roads in very visible and challenging pilot studies. Current research puts an emphasis on technology looking at aspects such as reliability and safety, and also on the interactions between the vehicles and other agents like human beings, conventional vehicles, and the existing network. Despite this rapid change, there are still many open questions regarding their application in the future, especially in what regards their use as public transport.

The contribution of this study is to provide a mathematical model to plan a last mile reservation based AT system to provide access to train stations. We considered two schemes: free (S1) and full service (S2). In the first, the company may select operational zones and the trips which it wants to satisfy and in the second when selecting a zone all trips must be satisfied. The models were both established with the objective to maximize the daily profit of such system by deciding on service zone locations and reservations to be accepted. Additionally, we allow the taxis to move without passengers between service zones. This is the potentially most significant benefit of AVs compared to conventional taxis: ATs can move without any driver which reduces operating costs. The models consider the vehicles to be electric thus some constraints were included in order to guarantee that on average taxis are idle enough time for the charging to happen.

The models were applied to the case study of Delft Zuid train station in Delft, the Netherlands. This case study provided interesting insights into the effect of service zone location and trip selection on the profitability of the AT system. From that application, we were able to take the following conclusions.

The first conclusion is that zone location and trip selection are able to diminish the negative impact of taxi imbalance hence contributing to the maximization of the profit. When the system is free to choose requests to serve (S1), the daily profit is always higher than the scheme with full service (S2), because this cannot reject inconvenient trips for the system without cancelling a whole operational zone. However model S2 has the advantage of fully covering a zone hence it guarantees a level of service at the expense of a lower profit. The second is that fleet size is an important factor on the profitability of the AT system because it influences the service capacity offered to passengers, as well as the total sunk cost of the fleet ownership. This conclusion can be drawn from observing the optimal daily profit of the two models where for 40 taxis under S1 and 60 taxis under S2 the maximum profit is achieved. The third conclusion is that when the model is free to choose trips (S1), having electric ATs constrains the system in accepting more requests for small fleets. For bigger fleets, the time spent at the train station should on average be enough to charge them. But the optimal results do not change using conventional taxis or electric ones in S2 since full service is more limitative than getting enough power. The fourth conclusion is that using ATs can significantly decrease the relocation costs and improve the daily profit of such taxi company

comparing to manpower relocation. In model S1, with a small fleet, this staff cost does not change the optimal solution but with bigger fleets, the optimum solutions point to serve fewer trips to avoid paying more for drivers, which leads to less revenue for the taxi company. For model S2, with bigger fleets, the optimal solution does not change, which indicates that full service has more influence than adding human driver costs in the objective function.

Despite the current version of the optimization models being practice-ready, there are some possible improvements which will make this model more realistic. First, we intend to take into account the negative impact of rejecting travelling requests by generating a penalty in the objective function. Moreover, in this chapter, we only apply the model under fixed total demand and price, but it would be interesting to assess the effects of varying prices and total demand on the daily profit of the system. Third, this model only considers a service between a train station and an urban area, however, there may be synergy effects from combining these trips with inter-zonal trips.

Chapter 3

Optimizing the dial-a-ride problem of ATs with dynamic travel times

In the previous chapter, ATs are used for the last mile problem. And the travel times between origin and destinations are assumed to be static. In this chapter, we propose a more general automated taxi service for full coverage of urban mobility. This system is envisioned to provide a transport service within a city area with a seamless door-to-door connection for all passengers' origins and destinations, much like what the existing transportation network companies (TNC) do. This problem defined as a DARP to design the routes and the schedules of these vehicles to serve passengers' requests. It also considers the effect on travel time by traffic congestion. This will be especially important when the number of automated vehicles circulating on the roads is so high that this will cause traffic congestion. We propose a linear integer programming (LIP) model to define the routing of the vehicles according to a profit maximization function while depending on dynamic travel times which vary with the flow of the ATs. The non-linear travel time function is linearized as discrete breakpoints to make the formulation solvable. The model is applied to a small road network.

This chapter is structured as follows: Section 3.1 introduces the existing research about ATs and the research gap for AT's DARP. Section 3.2 presents the mathematical model considering traffic congestion. Section 3.3 and section 3.4 apply the proposed model to a small case study and discuss the results gained from the model application. Section 3.5 draws conclusions from the model results.

This chapter is an edited version of the following paper:

Liang, X., Correia, G.H. de A., van Arem, B., 2017. An optimization model for vehicle routing of automated taxi trips with dynamic travel times. *Transportation Research Procedia* 27, 736–743. doi:10.1016/j.trpro.2017.12.045

3.1 Introduction

In recent years, technology development has accelerated the future roll-out of vehicle automation. This technology is expected to have a beneficial impact on travel efficiency, especially on interurban roads. Studies have mostly used micro-simulation tools or mathematical models to estimate the changes in road capacity and congestion under different levels of vehicle automation and cooperation. (Arem et al., 2006; Bose and Ioannou, 2003; Calvert et al., 2011).

The model that we propose in this chapter is related to the VRP, which is to design the best routes to provide services from a depot to some customers distributed in the network (Atasoy et al., 2015; Braekers et al., 2016; Figliozzi, 2011; Garcia-Najera and Bullinaria, 2010; Laporte, 2009; Spieser et al., 2014). It was first proposed by Dantzig and Ramser in 1959, as a generalization of the travel salesman problem. Traditionally, the vehicles start from the depot and then visit several customers that need to be served. The classic VRP consists in finding a set of routes for a fleet of identical vehicles, with the condition that each customer should be visited exactly once while minimizing the total routing cost (Pillac et al., 2013). In fact, the VRP can be seen as a class of problems since it has many variations based on the diversity of operating conditions and constraints when being applied in practice. Beyond the classical formulation, the capacitated VRP is the most common one, where vehicles have finite capacity when transporting passengers or goods. The VRP with time windows imposes that each customer must be visited during a specific time window (Mahmoudi and Zhou, 2016). When the original depot is not important anymore, and the system is asked to move freight loads or passengers from one node to another one in the road network, it becomes a pick-up and delivery problem. The most relevant variation to the VRP to this chapter is the DARP, which involves transporting people from their origins to their destinations (Braekers et al., 2014; Ho et al., 2018; Jaw et al., 1986) within requested time windows. According to the VRP variations that have been described, we may classify our model as the capacitated DARP with request time windows.

The problem that we are studying differs significantly from the previous research on DARP due to the large-scale network application. Generally, traditional VRP, including DARP, is focusing on tracking routes, which reflects the visiting sequences of each vehicle. In addition,

it assumes that the travel times between a fixed pair of nodes are deterministic regardless of the number of routing vehicles, even though those travel times may change along the day according to some pre-determined patterns. This assumption remains in most practical applications of ATs since the fleet is small compared with the number of other vehicles driving in the same network. In this chapter, we are considering an AT system with a larger number of vehicles, meaning that the dynamic traffic effect should not be ignored. To optimize the AT's routing under congestion, we introduce the formulation of traffic assignment to the road network integrated with the DARP.

Traffic assignment is an important component of the classic four-stage transport model, which is defined as to assign the trips by each mode to their corresponding networks (Dios Ortúzar and Willumsen, 2000). For car traffic, it determines the traffic flows on each link in a road network while involving a solution to the equilibrium between the supply side (the network system) and the demand side (the road users). Traffic assignment takes into account the congestion effects of the route choices of private drivers which is useful to be incorporated within the DARP in order to test the large scale network application of an AT system.

A key step in the traffic assignment is to load a trip matrix (demand) onto the network and produce a set of link flows. During the loading process, different route choice conditions of the drivers may require different load methods. In the traditional traffic assignment based on human driving, different individuals may perceive the best route in different ways: some would always prefer the uninterrupted flow conditions; while others would like to use the shortest distance routes. This effect is referred to as the stochastic elements in the route choice of the traffic assignment. In the automated driving circumstance, the stochastic effects are eliminated since the system gives the control of the vehicle, thus its route, to a computer making all ATs aware of the best route under the same performance indicators. The absence of the stochastic effects is not an assumption of this research to simplify the traffic assignment model, but an important feature of the AT system compared to human-driven cars.

The distribution of the flows of conventional vehicles in a network follows the well-known traffic engineering principle of user equilibrium as formally proposed by Wardrop: "Under equilibrium conditions traffic arranges itself in congested networks in such a way that no individual trip maker can reduce his path costs by switching routes (Wardrop, 1952)". The user equilibrium principle was applied in the routing of private AVs in a city in the already mentioned study from Correia and van Arem (2016), where the authors hypothesized that households would have control of the routes of their private cars and could act selfishly by minimizing their own travel costs.

Wardrop also proposed the second principle of assigning car traffic to a road network, which is referred to as system equilibrium or system optimum: "Under social equilibrium conditions, traffic should be arranged in congested networks in such a way that the average (or total) travel cost is minimized (Wardrop, 1952)." The user equilibrium attempts to model the behaviour of individual drivers with minimizing their own travel costs, while the social equilibrium is a design principle oriented towards the system planners, trying to manage the

traffic on the road network and minimize the total costs of all the travellers (Peeta and Mahmassani, 1995). The system optimum can only happen when all drivers agree with the set of paths to be chosen or these paths are enforced by an externality, which is corresponding to the nature of an AT system.

A mathematical programming framework proposed in Beckmann et al. (1955) is commonly used as a congested assignment technique to approximate Wardrop's equilibrium. The formulation is a non-linear programming problem due to the functions which represent the traffic flow and capture the complex interactions among vehicles. This can be solved in two main ways: linearizing the problem through simplifications (An and Lo, 2015) or using iterative algorithms (Correia and van Arem, 2016a).

An iterative algorithm starts with an initial set of link costs and link flows. During each iteration, the current traffic flow is a linear function of the previous iteration's flow and the present flow resulting from an all-or-nothing assignment. The iteration progress stops when the flows do not change significantly according to the criteria set up. Correia and van Arem (2016) used the same principle to tackle the incremental process of assigning each household private automated vehicle to the road network. The method avoids the non-linearity of the cost-flow function inside the mathematical program to route the vehicles while they pick up and drop off elements of the family. Instead, the travel times resulting from the Bureau of Public Roads (BPR) function are updated after the assignment process of all the vehicles in each iteration. In a later work, Levin (2017) also considered the traffic congestion effects when studying the routing of a large number of shared AVs by using a link-transmission model. The model was applied to a small network due to the considerable computation time of the method. These two papers propose methods to route private or shared AVs when considering the influence of traffic congestion caused by the vehicles themselves. However, they both need the demand to be known before the solving process, which does not fit a system with real-time requests.

In this chapter, we are going to study a transport system with a fleet of automated vehicles used as taxis which provide a transport service within a city area with a seamless door-to-door connection for passengers' origins and destinations. Then a model is proposed to address AT's DARP, meaning that assigning the vehicles to paths on an urban road network while satisfying the demand. Moreover, this model is expected to solve the problem with dynamic travel times which vary with the flow of the ATs. We linearize the travel time-traffic flow function as a way to solve the non-linearity of Beckmann's formulation, which makes it possible to achieve the optimal distribution of vehicles as the result of the mathematical programming model.

3.2 An integer programming model with discrete travel time function

In this section, we describe a LIP model whose objective is to determine the optimal vehicle routing of the AT system.

The notations used in this chapter are presented in Table 3.1.

Table 3.1 Notations

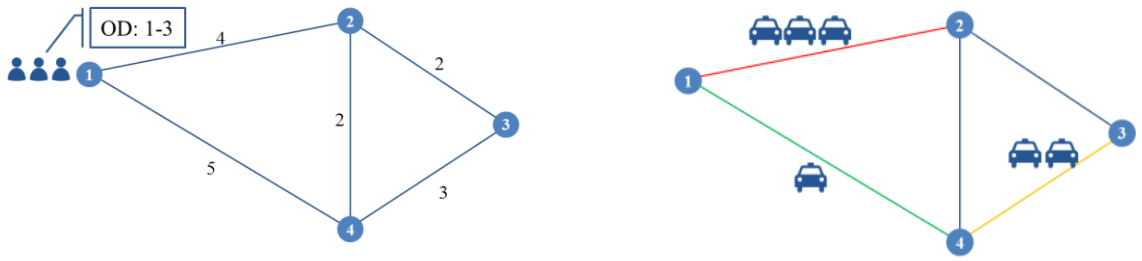
Notation	Description
Sets	
$\boldsymbol{I}$	$= \{1, \dots, i, \dots I\}$, set of the nodes in the network, where N is the total number of nodes.
$\boldsymbol{T}$	$= \{0, \dots, t, \dots, T\}$, set of time instants in the service period, where T is the total number of time steps in the service period.
$\boldsymbol{E}$	$= \{1, \dots, e, \dots, E\}$, set of trips, where E is the total number of all the trips in the service period.
$\boldsymbol{V}$	$= \{1, \dots, v, \dots, V\}$, set of vehicles, where V is the total number of taxis in the system.
$\boldsymbol{K}$	$= \{0, \dots, k, \dots, K\}$, set of breakpoints for dynamic travel time, where K is the total number of the breakpoints in the flow-travel time curve.
$\boldsymbol{G}$	$= \{\dots, (i,j), \dots\}$, set of links of the road network where vehicles move, $i, j \in \boldsymbol{I}$, $i \neq j$.
$\boldsymbol{M}$	$= \{\dots, (i_{t_1}, j_{t_2}), \dots\}$, set of links in the time-space network, $\forall (i,j) \in \boldsymbol{G}, \forall t_1, t_2 \in \boldsymbol{T}, t_1 < t_2, \delta_{ij}^{min} \leq t_2 - t_1 \leq \delta_{ij}^{max}$.
Data	
a^e	desired departure time for the eth trip, $e \in \boldsymbol{E}$.
b^e	latest arrival time for the eth trip, $e \in \boldsymbol{E}$.
c^e	earliest departure time for the eth trip, $e \in \boldsymbol{E}$.
w^e	maximum waiting time for the eth trip, $e \in \boldsymbol{E}$.
ori^e	the origin node for the eth trip, $e \in \boldsymbol{E}$.
des^e	the destination node for the eth trip, $e \in \boldsymbol{E}$.
δ_{ij}^{max}	maximum travel time from node i to node j, $\forall (i,j) \in \boldsymbol{G}$.
δ_{ij}^{min}	minimum travel time from node i to node j, $\forall (i,j) \in \boldsymbol{G}$.
δ_{ij}^{k}	travel time from node i to node j at breakpoint k, $\forall (i,j) \in \boldsymbol{G}, k \in \boldsymbol{K}$.
d_{ij}	travel distance from node i to node j, $\forall (i,j) \in \boldsymbol{G}$.
opt_{ij}	shortest travel time from node i to node j, $\forall i, j \in \boldsymbol{I}$ (computed with δ_{ij}^{min}).
$rcap_{ij}$	Capacity of each link (i,j) which is the number of vehicles that go through the link at the highest travel time, $\forall (i,j) \in \boldsymbol{G}$.

Parameters	
c_r	price rate of taxi service, euro/time-step.
c_f	fuel cost per driving distance, euro / km.
c_q	parking cost, euro/time step.
c_v	vehicle depreciation cost per vehicle, euro/vehicle
c_p	penalty cost per trip if a request is rejected, euro/ request.
c_w	penalty for passengers' waiting, euro/time step.
c_d	penalty for delivery delay, where the delay is defined as the real travel time minus the shortest travel time for a request, euro/time step.
Decision variables	
$x^v_{i_{t_1},j_{t_2}}$	binary variable equals to 1 if vehicle v drives from node i to j from time instant t, $\forall(i_{t_1},j_{t_2}) \in \boldsymbol{M}, \forall v \in \boldsymbol{V}$.
$y^v_{i_t}$	binary variable equals to 1 if vehicle v parks at i from time instant t to $t+1$, $\forall i \in \boldsymbol{I}$, $\forall v \in \boldsymbol{V}$, $\forall t \in \boldsymbol{T}, t < T$.
S^{ev}_{ij}	binary variable equals to 1 if trip e from node $i = ori^e$ to node $j = des^e$ is done using vehicle v, $\forall e \in \boldsymbol{E}$, $\forall v \in \boldsymbol{V}$.
P^{evt}_{ij}	binary variable equals to 1 if trip e from node $i = ori^e$ to node $j = des^e$ done by vehicle v departs at time instant t, $\forall e \in \boldsymbol{E}$, $\forall v \in \boldsymbol{V}$, $\forall t \in \boldsymbol{T}, c^e \leq t \leq a^e + w^e$.
A^{evt}_{ij}	binary variable equals to 1 if trip e from node $i = ori^e$ to node $j = des^e$ *done* by vehicle v arrives at time instant t, $\forall e \in \boldsymbol{E}$, $\forall v \in \boldsymbol{V}$, $\forall t \in \boldsymbol{T}, a^e + opt_{ij} \leq t \leq b^e$.
ϕ^{ev}_{ij}	integer variable of the difference between the desired pick-up time and real pick-up time of trip e from node $i = ori^e$ to node $j = des^e$ done by vehicle v, $\forall e \in \boldsymbol{E}$, $\forall v \in \boldsymbol{V}$.
L^v_t	binary variable equals to 1 if vehicle v is transporting a passenger from time instant t to $t+1$, $\forall v \in \boldsymbol{V}$, $\forall t \in \boldsymbol{T}', t < T$.
F^t_{ij}	flow of vehicles on link (i,j) starting from time instant t, $\forall(i,j) \in \boldsymbol{G}$, $\forall t \in \boldsymbol{T}, t < T$.
$\lambda^{t,k}_{ij}$	binary variable equals to 1 if the traffic flow on link (i,j) starting from t is taking the value on breakpoint k of the flow-travel time curve, $\forall(i,j) \in \boldsymbol{G}$, $\forall t \in \boldsymbol{T}, t < T$, $\forall k \in \boldsymbol{K}$.
δ^t_{ij}	travel time when travelling on link (i,j) starting from time instant t, $\forall(i,j) \in \boldsymbol{G}$, $\forall t \in \boldsymbol{T}, t < T$.

3.2.1 System setting

In Figure 3.1 we show a small example to illustrate how this system works. The values on each link show the shortest travel time (free flow travel time) for that link. In this case, two passengers plan to travel from node 1 to node 3 (Figure 3.1 (a)). For these two passengers, the shortest path would be 1-2-3 with travel time 6 in total. However, the shortest travel time is

not applicable all the time since the travel time will be influenced by how many ATs are travelling on that link. When traffic congestion happens on link 1-2 of the shortest path 1-2-3 (Figure 3.1 (b)), the travel time will increase, which makes other possible paths like 1-4-3 competitive or even shorter than path 1-2-3. Therefore, an optimization model is needed to decide which paths the ATs should choose to satisfy clients' travel requests based on the dynamic travel time, which is defined as a function of traffic flow.



(a) Network with travel requests **(b) Network with dynamic travel times**

Figure 3.1 An example network with travel requests and dynamic travel times

Before presenting the proposed model, it is important to state all the assumptions which were considered in the formulations. We consider a future scenario in which the ATs will replace all modes of personal transport thus the alternative travel modes will be mass public transport e.g. metro, bus and tram. Since these alternative transport modes are usually seen as not contributing to the congestion in the network, we do not consider background traffic flow for simplification meaning that the flow is generated only by the ATs themselves. This model solves a static DARP problem, which needs the demand to be known in advance. This AT system provides service to individual trips which means ride-sharing with strangers is not considered. We consider a private taxi service between any pair of nodes within the city road network, with no background traffic flow which means the flow is generated only by the ATs themselves. The model works on the assumption that the taxi company can achieve total control of the system by being free to accept or reject requests according to a profit maximization objective. If the travel request cannot be fulfilled, then a penalty will occur: we assume that the AT company would have to pay compensation which can be viewed as paying partially the cost of an alternative transport like metro or bus.

In this chapter, we need the demand for the whole optimization period pre-known as the system input. Passengers use an online app to book an AT providing travel information that includes the origin, the destination and the time they would like to departure. Figure 3.2 shows the relation of time components for each request. The leading time is the measure indicating how far in advance the requests are claimed. If the leading times for all the requests are long enough, the system will receive perfect information before ahead. This is a static optimization problem. Waiting for late departure is allowed. Since the traffic condition is

considered, the real travel time for each request may be longer than the shortest travel time. We set a limitation for the longest travel time in order to guarantee service quality.

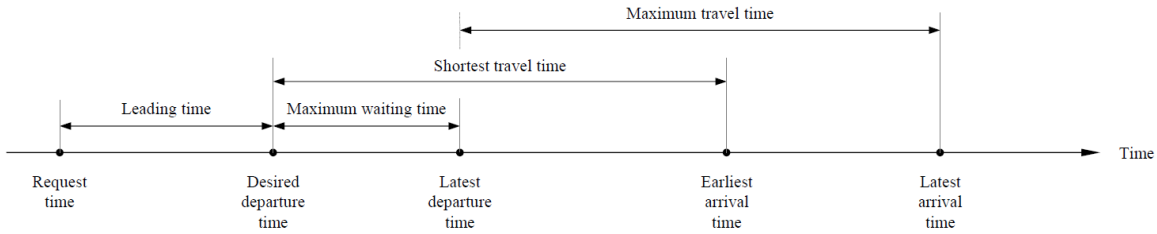


Figure 3.2 Time components for each request

3.2.2 A linear integer programming formulation

The [$\boldsymbol{LIP}$] model for solving the problem defined above has the following formulation.

Objective function:

$$[\boldsymbol{LIP}] \quad \max Z = Pr \cdot \sum_{\substack{e\in \boldsymbol{E}, v\in \boldsymbol{V} \\ i=ori^e \\ j=des^e}} S_{ij}^{ev} \cdot opt_{ij} - c_f \cdot \sum_{\substack{(i_{t_1}, j_{t_2})\in \boldsymbol{M}, \\ v\in \boldsymbol{V}}} x_{i_{t_1}, j_{t_2}}^{v} \cdot d_{ij} - c_v \cdot V - c_q$$

$$\cdot \sum_{\substack{i\in \boldsymbol{I}, v\in \boldsymbol{V} \\ t\in \boldsymbol{T}, t<T}} y_{i_t}^{v} - c_p \cdot \sum_{e\epsilon \boldsymbol{E}} \left(1 - \sum_{\substack{v\in \boldsymbol{V} \\ i=ori^e \\ j=des^e}} S_{ij}^{ev} \right) - C_w \cdot \sum_{\substack{e\epsilon \boldsymbol{E}, v\in \boldsymbol{V} \\ i=ori^e \\ j=des^e}} \phi_{ij}^{ev} - C_d \tag{3.1}$$

$$\cdot \sum_{\substack{e\epsilon \boldsymbol{E}, v\in \boldsymbol{V} \\ i=ori^e \\ j=des^e}} \left(\sum_{t\in \boldsymbol{T}} (A_{ij}^{evt} \cdot t) - \sum_{t\in \boldsymbol{T}} (P_{ij}^{evt} \cdot t) - opt_{ij} \cdot S_{ij}^{ev} \right)$$

The objective function (3.1) is considered from both the AT company and the passengers' perspective. It is a generalized cost-benefit summation for the two components of the system. For the AT company, it aims to maximize the total profit which includes the total revenue from the customers, the vehicle fuel costs, the vehicle depreciation costs and the parking costs. The revenue is only charged based on the shortest path of each request, if that was not the case the model would try to detour passengers to charge them for more travel distance and time. The vehicle depreciation cost is considered based on the number of vehicles, which means that this cost is a constant component in the objective function and will not influence the solution space of the problem. However, we decide to keep the depreciation cost since we want to analyse the monetary impact of the fleet size on the system profit. Regarding the passengers, when a request is rejected, passengers are assumed to use public transport as an alternative to finish their trip and the AT system should pay compensation for the users as a

penalty. Late arrival and delivery delay are also penalized in order to consider the level of service offered to the clients.

Constraints:

$$S_{ij}^{ev} \leq \sum_{\substack{i_{t_1}, l_{t_2} \epsilon \boldsymbol{M} \\ t_1 \geq c^e, t_2 \leq b^e}} x_{i_{t_1}, l_{t_2}}^{v} \qquad \forall e \in \boldsymbol{E}, v \in \boldsymbol{V}, i = ori^e, j = des^e \tag{3.2}$$

Constraints (3.2) describe that trip e from node i to node j can only be satisfied by vehicle v if that vehicle has pass through node i between the earliest departure time c^e and the latest arrival time b^e.

$$S_{ij}^{ev} \leq \sum_{\substack{l_{t_1}, j_{t_2} \epsilon \boldsymbol{M} \\ t_1 \geq c^e, t_2 \leq b^e}} x_{l_{t_1}, j_{t_2}}^{v} \qquad \forall v \in \boldsymbol{V}, e \in \boldsymbol{E}, i = ori^e, j = des^e \tag{3.3}$$

Constraints (3.3) describe that trip e from node i to node j can only be satisfied by vehicle v if that vehicle has pass through node j between the earliest departure time c^e and the latest arrival time b^e.

$$\sum_{\substack{t \epsilon \boldsymbol{T} \\ c^e \leq t \leq a^e + w^e}} P_{ij}^{evt} = S_{ij}^{ev} \qquad \forall e \in \boldsymbol{E}, v \in \boldsymbol{V}, i = ori^e, j = des^e \tag{3.4}$$

Constraints (3.4) assure that a satisfied trip must have only one departure time or this trip is not satisfied at all.

$$\sum_{\substack{t \epsilon \boldsymbol{T} \\ c^e + opt_{ij} \leq t \leq b^e}} A_{ij}^{evt} = S_{ij}^{ev} \qquad e \in \boldsymbol{E}, v \in \boldsymbol{V}, i = ori^e, j = des^e \tag{3.5}$$

Constraints (3.5) assure that a satisfied trip must have only one arrival time or this trip is not satisfied at all.

$$\sum_{\substack{t \epsilon \boldsymbol{T} \\ c^e \leq t \leq a^e + w^e}} P_{ij}^{evt} \leq 1 \qquad e \in \boldsymbol{E}, v \in \boldsymbol{V}, i = ori^e, j = des^e \tag{3.6}$$

Constraints (3.6) guarantee that one trip can only have one departure time when it is served or this trip is not satisfied by any vehicle.

$$\sum_{\substack{t \in \boldsymbol{T} \\ c^e + opt_{ij} \leq t \leq b^e}} A_{ij}^{evt} \leq 1 \qquad e \in \boldsymbol{E}, v \in \boldsymbol{V}, i = ori^e, j = des^e \tag{3.7}$$

Constraints (3.7) guarantee that one trip can only have one arrival time when it is served or this trip is not satisfied by any vehicle.

$$\sum_{\substack{t\in \boldsymbol{T} \\ c^e\le t\le a^e+w^e}} (P_{ij}^{evt}\cdot t) \le \sum_{\substack{t\in \boldsymbol{T} \\ c^e+opt_{ij}\le\, t\le b^e}} \left(A_{ij}^{evt}\cdot t\right) \quad e\in \boldsymbol{E}, v\in \boldsymbol{V}, i=ori^e, j=des^e \tag{3.8}$$

Constraints (3.8) impose that the departure time instant of a satisfied trip must happen before the arrival time instant.

$$\phi_{ij}^{ev} = \sum_{t\in \boldsymbol{T}} P_{ij}^{evt}\cdot t - a^e\cdot S_{ij}^{ev} \quad \forall e\in \boldsymbol{E}, v\in \boldsymbol{V}, i=ori^e, j=des^e \tag{3.9}$$

Constraints (3.9) computes the difference between the desired departure time and the real departure time which is the passengers' waiting time.

$$\sum_{v\in \boldsymbol{V}} S_{ij}^{ev} \le 1 \quad \forall e\in \boldsymbol{E}, i=ori^e, j=des^e \tag{3.10}$$

Constraints (3.10) ensure that a trip could only be served by one vehicle.

$$\sum_{i\in \boldsymbol{I}} y_{i_t}^{v} \le 1 \quad \forall v\in \boldsymbol{V}, t\in \boldsymbol{T}, t<T \tag{3.11}$$

Constraints (3.11) impose that for a specific time step $t\ to\ t+1$ one automated taxi should only park at one place node or not park at all.

$$\sum_{i_{t_1},j_{t_2}\in \boldsymbol{M}} x_{i_{t_1},j_{t_2}}^{v} \le 1 \quad \forall i\in \boldsymbol{I}, v\in \boldsymbol{V}, t_1\in \boldsymbol{T} \tag{3.12}$$

Constraints (3.12) guarantee that one trip with specific origin and departure time can only have one destination and one arrival time.

$$\sum_{\substack{i_{t_1},j_{t_2}\in \boldsymbol{M} \\ t_1\le t,\, t_2>t}} x_{i_{t_1},j_{t_2}}^{v} + \sum_{i\in \boldsymbol{I}} y_{i_t}^{v} = 1 \quad \forall v\in \boldsymbol{V}, t\in \boldsymbol{T}, t<T \tag{3.13}$$

Constraints (3.13) ensure that each vehicle can only have one status at time instant t: either going to the next destination or parking at that place.

$$\sum_{i_t,j_{t_1}\in \boldsymbol{M}} x_{i_t,j_{t_1}}^{v} + y_{i_t}^{v} = \sum_{i_{t_2},j_t\in \boldsymbol{M}} x_{j_{t_2},i_t}^{v} + y_{i_{t-1}}^{v} \quad \forall i\in \boldsymbol{I}, t\in \boldsymbol{T}, t<T, v\in \boldsymbol{V} \tag{3.14}$$

Constraints (3.14) are flow conservation constraints which make sure that the number of taxis leaving from node i and parking there from time instant t is equal to the number of vehicles arriving at node i and parking at the same place until t.

$$\sum_{\substack{i_0, j_{t_1} \in \boldsymbol{M} \\ v \in \boldsymbol{V}}} x_{i_0, j_{t_1}}^{v} + \sum_{v \in \boldsymbol{V}} y_{i_0}^{v} = V \qquad i = initial\ parking\ station \tag{3.15}$$

Constraints (3.15) describe the initial status of the AT fleet. At the beginning of the operation period, all vehicles are stocked at the initial parking station.

$$F_{ij}^{t_1} = \sum_{v \in \boldsymbol{V}} x_{i_{t_1}, j_{t_2}}^{v} \qquad \forall (i_{t_1}, j_{t_2}) \in \boldsymbol{M} \tag{3.16}$$

Constraints (3.16) compute the flow of vehicles in each road link (i, j) from time instant t_1.

$$F_{ij}^{t} \leq rcap_{ij} \quad \forall (i, j) \in \boldsymbol{G}, \forall t \in \boldsymbol{T}, t < T \tag{3.17}$$

Constraints (3.17) limit flow by the capacity of each link.

In this chapter, we consider the travel time as a function of the traffic flow by the Bureau of Public Roads (Dafermos and Sparrow, 1969): $t(V) = t_0(1 + a \times (\frac{V}{Q})^b)$, where $t(V)$ is the travel time when the traffic flow is V, t_0 is the zero-flow travel time; V is the traffic flow; Q is the road capacity; a and b are estimation parameters (shown in Figure 3.3 (a)). Using this function directly to build the constraints will make the model non-linear, which is difficult to solve. Therefore, we introduce a parameter k to indicate the congestion level on each road link and obtain the dynamic travel time by discretizing the traffic flow into several breakpoints (shown in Figure 3.3 (b)). Using this method, the non-linear travel time function is linearized and then the proposed problem is solvable.

$$F_{ij}^{t} = \sum_{k \in \boldsymbol{K}} \lambda_{ij}^{t,k} \cdot k \qquad \forall (i, j) \in \boldsymbol{G}, \forall t \in \boldsymbol{T}, t < T \tag{3.18}$$

$$\sum_{k \in \boldsymbol{K}} \lambda_{ij}^{t,k} = 1 \qquad \forall (i, j) \in \boldsymbol{G}, \forall t \in \boldsymbol{T}, t < T \tag{3.19}$$

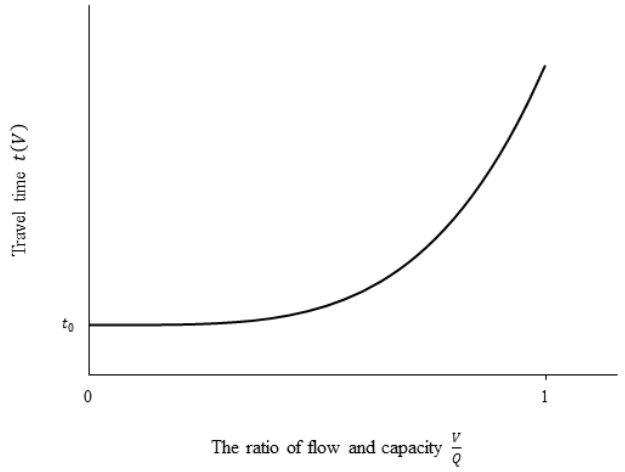
Constraints (3.18)-(3.19) impose a break-point k to describe the congestion level of each link.

$$\delta_{ij}^{t} = \sum_{k \in \boldsymbol{K}} \delta_{ij}^{k} \cdot \lambda_{ij}^{t,k} \qquad \forall (i, j) \in \boldsymbol{G}, \forall t \in \boldsymbol{T}, t < T \tag{3.20}$$

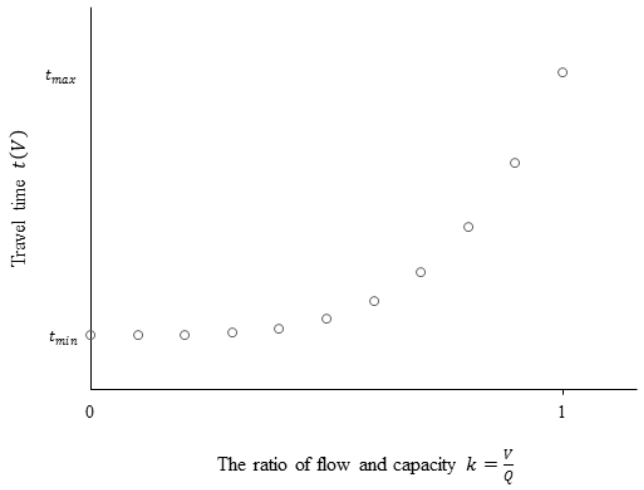
Constraints (3.20) define the travel time on road link (i, j) from time instant t related to the congestion level represented by the breakpoint k.

$$\delta_{ij}^{t_1} \leq (t_2 - t_1) \cdot x_{i_{t_1}, j_{t_2}}^{v} + t_{ij}^{max} \cdot \left(1 - x_{i_{t_1}, j_{t_2}}^{v}\right) \qquad \forall (i_{t_1}, j_{t_2}) \in \boldsymbol{M}, v \in \boldsymbol{V} \tag{3.21}$$

$$\delta_{ij}^{t_1} \geq (t_2 - t_1) \cdot x_{i_{t_1}, j_{t_2}}^{v} + t_{ij}^{min} \cdot \left(1 - x_{i_{t_1}, j_{t_2}}^{v}\right) \qquad \forall (i_{t_1}, j_{t_2}) \in \boldsymbol{M}, v \in \boldsymbol{V} \tag{3.22}$$



(a) classic BPR function



(b) travel time function based on discretized flow

Figure 3.3 The travel time function

Constraints (3.21)-(3.22) impose the travel time is not greater than the maximum travel time or smaller than the minimum travel time. Meanwhile, they also guarantee that when vehicle v travelling on link $(i_{t_1} j_{t_2})$, the travel time must be $\delta_{ij}^{t_1}$.

$$t_1 + \delta_{ij}^{t_1} \leq t_2 + \delta_{ij}^{t_2} \qquad \forall (i,j) \in \boldsymbol{G}, t_1, t_2 \in \boldsymbol{T}, t_1, t_2 < T, t_1 < t_2 \tag{3.23}$$

Constraints (3.23) guarantee link consistency. We build the assumption that vehicles entering in a link (i,j) later from node i should not leave earlier from node j, which means the ATs do not pass on another. These constraints were established by Kaufman et al.(1992).

$$L_t^v = \sum_{\substack{i=ori^e, j=des^e \\ e\in \boldsymbol{E}, \forall t_1 \in \boldsymbol{T}, t_1 \le t}} P_{ij}^{evt_1} - \sum_{\substack{i=ori^e, j=des^e \\ e\in \boldsymbol{E}, \forall t_2 \in \boldsymbol{T}, t_2 \le t}} A_{ij}^{evt_2} \qquad \forall t \in \boldsymbol{T}, t < T, v \in \boldsymbol{V} \tag{3.24}$$

Constraints (3.24) computes the number of passengers loaded on each vehicle during time step t to $t+1$.

$$L_t^v + \sum_{i\in \boldsymbol{I}} y_{i_t}^v \le 1 \qquad \forall t \in \boldsymbol{T}, t < T, v \in \boldsymbol{V} \tag{3.25}$$

Constraints (3.25) force that the ride-sharing or parking with passengers occupied is not allowed.

$$x_{i_{t_1}, j_{t_2}}^v \in \{0,1\} \qquad \forall (i_{t_1}, j_{t_2}) \in \boldsymbol{M}, \forall v \in \boldsymbol{V} \tag{3.26}$$

$$y_{i_t}^v \in \{0,1\} \qquad \forall i \in \boldsymbol{I}, \forall v \in \boldsymbol{V}, \forall t \in \boldsymbol{T}, t < T \tag{3.27}$$

$$S_{ij}^{ev} \in \{0,1\} \qquad i = ori^e, j = des^e, \ \forall e \in \boldsymbol{E}, \forall v \in \boldsymbol{V} \tag{3.28}$$

$$P_{ij}^{evt} \in \{0,1\} \qquad i = ori^e, j = des^e, \ \forall e \in \boldsymbol{E}, \forall v \in \boldsymbol{V}, \forall t \in \boldsymbol{T}, c^e \le t \le a^e + w^e \tag{3.29}$$

$$A_{ij}^{evt} \in \{0,1\} \qquad i = ori^e, j = des^e, \forall e \in \boldsymbol{E}, \forall v \in \boldsymbol{V}, \forall t \in \boldsymbol{T}, a^e + opt_{ij} \le \ t \le b^e \tag{3.30}$$

$$\phi_{ij}^{ev} \in \boldsymbol{N^0} \qquad i = ori^e, j = des^e, \ \forall e \in \boldsymbol{E}, \forall v \in \boldsymbol{V} \tag{3.31}$$

$$L_t^v \in \{0,1\} \qquad \forall v \in \boldsymbol{V}, \forall t \in \boldsymbol{T}, t < T \tag{3.32}$$

$$F_{ij}^t \in \boldsymbol{N^0} \qquad \forall (i,j) \in \boldsymbol{G}, \forall t \in \boldsymbol{T}, t < T \tag{3.33}$$

$$\lambda_{ij}^{t,k} \in \{0,1\} \qquad \forall (i,j) \in \boldsymbol{G}, \forall t \in \boldsymbol{T}, t < T, \forall k \in \boldsymbol{K} \tag{3.34}$$

$$\delta_{ij}^t \in \boldsymbol{N^0} \qquad \forall (i,j) \in \boldsymbol{G}, \forall t \in \boldsymbol{T}, t < T \tag{3.35}$$

Constraints (3.26)-(3.35) define the domain for the decision variables.

3.2.3 Valid inequalities

The searching process for the optimal solution of the AT routing and assignment problem **[*LIP*]** can be accelerated by bounding the problem with extra constraints. They do not change the feasible solution region but tight the bounds on that space by eliminating non-integer solutions of the relaxation process of the traditional branch and bound search method. Thus the gap to find the optimal solution closes faster since certain nodes of the tree will not be explored.

$$\sum_{j\in \boldsymbol{I}, t_1 \in \boldsymbol{T}} x_{i_t, j_{t_1}}^v + y_{i_t}^v \le 1 \quad \forall i \in \boldsymbol{I}, \forall v \in \boldsymbol{V}, \forall t \in \boldsymbol{T}, t < T, \tag{3.36}$$

Constraints (3.36) guarantee that if vehicle v is at node i at time instant t, it can only have one status: either going to the next destination or parking at node i; otherwise vehicle v is at somewhere else. These constraints are already achieved by constraints (3.12) and (3.13), where constraints (3.12) initiate the existence of vehicle v and constraints (3.13) maintain its existence throughout the whole optimization period. In the branch and bound process, it may happen that the relaxed variables on the left-hand side of constraints (3.13) sum to the same value of the right-hand side, where each variable obeying its domain constraints and the sums being above one. Imposing bound constraints will not allow those solution nodes to be explored in the relaxation process.

$$P_{ij}^{evt} \leq S_{ij}^{ev} \quad \forall e \in \boldsymbol{E}, v \in \boldsymbol{V}, i = ori^e, j = des^e, \forall t \in \boldsymbol{T}, c^e \leq t \leq a^e + w^e \tag{3.37}$$

$$A_{ij}^{evt} \leq S_{ij}^{ev} \quad \forall e \in \boldsymbol{E}, v \in \boldsymbol{V}, i = ori^e, j = des^e, \forall t \in \boldsymbol{T}, c^e + opt_{ij} \leq t \leq b^e \tag{3.38}$$

Constraints (3.37)-(3.38) impose that once a specific request is served, it can have a departure time and arrival time.

$$P_{ij}^{evt} \leq \sum_{\substack{i_t, l_{t_1} \epsilon \boldsymbol{M} \\ t_1 \leq b^e}} x_{i_t, l_{t_1}}^{v} \quad \forall e \in \boldsymbol{E}, v \in \boldsymbol{V}, t \in \boldsymbol{T}, i = ori^e, j = des^e \tag{3.39}$$

Constraints (3.39) ensure that trip e from node i to node j can only have departure time t by vehicle v if that vehicle has pass through node i at time instant t.

$$A_{ij}^{evt} \leq \sum_{\substack{l_{t_1}, j_t \epsilon \boldsymbol{M} \\ t_1 \geq c^e}} x_{l_{t_1}, j_t}^{v} \quad \forall e \in \boldsymbol{E}, v \in \boldsymbol{V}, t \in \boldsymbol{T}, i = ori^e, j = des^e \tag{3.40}$$

Constraints (3.40) ensure that trip e can only have arrival time t if vehicle v has pass through node j at time instant t. These constraints are already achieved by constraints (3.2)-(3.7), but imposing extra constraints (3.37)-(3.40) will avoid exploring the certain nodes and bound the solution region, which can accelerate the searching process.

3.3 Case study

In this section, we apply the model to a small toy network to illustrate how it works and the results that can be obtained. We consider a road network with 9 nodes and 12 links (Figure 3.4). The length of each link is 1 km and the central parking station is located at node 5. The capacity of each link is 3 vehicles per direction per time step. The minimum travel time (δ_{ij}^{min}) on each link is considered as 1 time step and the assumed maximum speed is 30km/h. With a simplified volume-delay function, the travel time values are 1, 2 and 4 time step respectively, when the traffic flows are 1, 2 and 3 vehicles respectively (Figure 3.5). When ATs parking at the central node 5, it is free to charge while parking at other nodes should be paid.

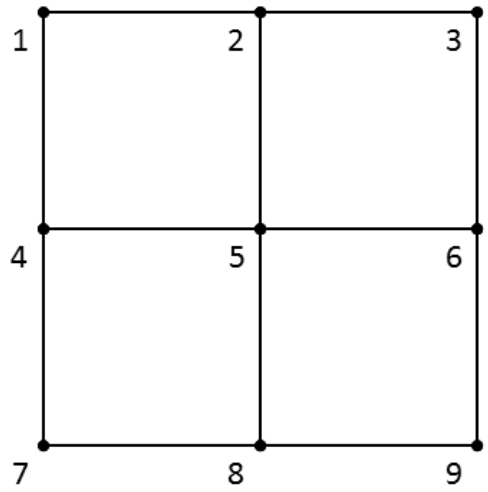


Figure 3.4: Road network with 9 nodes and 12 links

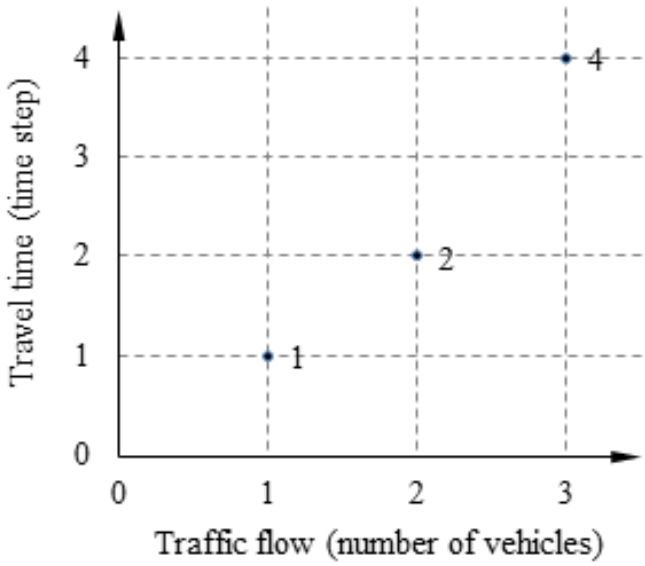


Figure 3.5: Traffic flow and travel time value

There are 10 time steps from time instant 0 to 10 in the service period. Each time step of the optimization is 2 minutes. A short time period with 2 time steps is set in order to be possible to satisfy the early trips and late trips in the service period, which makes the operation time run from -2 to 12 time instant. Using Monte Carlo simulation, we generate 20 requests each of which has an origin, a destination and a desired departure time based on an uneven distribution with a higher probability for node 5 to be an origin. In this case study, the number of ATs serving in the system is 5.

For this case study, we consider the following parameters: price rate c_r=2 euro/time-step vehicle fuel cost c_f=0.1 euro/km, vehicle depreciation cost c_v=5 euro/vehicle, waiting cost c_w

=0.05 euro/time step, penalty cost for using public transport c_p=2 euro/request, congestion delay penalty c_d=0.05 euro/time-step.

3.4 Experiments and results

Firstly a scenario with the parking cost c_q=0.05 euro/time step is tested and the graphical results of every vehicle's movements ($x^v_{i_{t_1},j_{t_2}}$) in the optimal solution are shown in Figure 3.6. In this figure, each solid line represents the AT's movement when it is satisfying passengers' demand while the dashed line represents an empty AT. Each vehicle's position at a specific time instant is shown by the number with a circle in this figure. Before the service period the ATs, are allowed to travel from central node 5 to any node in the network if it is needed. In this case, two ATs (vehicle 1and 4) are relocated before time instant 0 in order to pick up passengers in time.

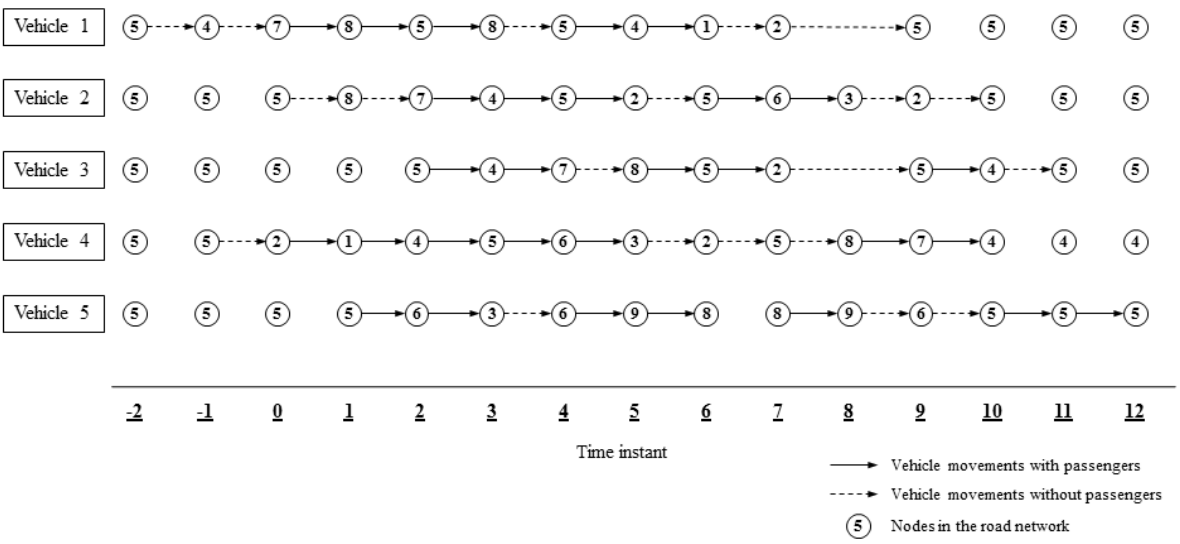


Figure 3.6 AT movement results

According to constraints (3.16)-(3.19), when more than one AT is travelling on the same link starting at the same time instant, traffic congestion must be formed. This can be seen in Figure 3.6. When only assigning vehicle 1 to road link (2,5) starting from time instant 1, the travel time would be one time step. Then vehicle 3 is also assigned to the same link at the same departure time hence this leads to a more crowded road link. Therefore the travel time for both vehicles is delayed.

A comparison between AT routing optimization with dynamic and static travel time was done and Table 3.2 presents the AT service results under different parking price rate scenarios. The column "Total idle time" represents the parking time at all the nodes in the road network while column "Paid parking time" is the total parking time at all the nodes except node 5. "Congestion delay" is computed as the total number of time steps between the real travel time and the shortest travel time for all the AT trips while "Number of delay requests" is only for

the trips when ATs are serving requests. The computational time is within one minute for all the scenarios and the static travel time scenario has a shorter computational time than the dynamic one when the parking price is the same.

Table 3.2 Optimization model results

Type of Travel time	Total profit (euro)	Parking price (euro/time step)	Number of satisfied requests (out of 20)	Total idle time (time step)	Paid parking time (time step)	% of congestion	Congestion delay (time step)	Total travel distance (km)	Computational time (s)
Dynamic	15.00	0.00	14	14	12	0%	0	40	34.4
Static	15.00		14	13	10	8%	-	40	1.3
Dynamic	14.15	0.05	15	5	1	2%	0	47	23.9
Static	14.15		15	11	7	10%	-	44	4.5
Dynamic	14.00	0.10	15	5	2	2%	0	40	12.6
Static	14.10		15	7	1	12%	-	40	5.4

When parking at any node of the road network is free, there is no difference in the optimal solution between the scenario with dynamic and static travel time. This is because the system is free to park anywhere at any time to avoid traffic congestion so more than one AT travelling on the same link is not needed. When travel time is static, the system is free to generate congestion on each link which brings 8% congested links in this scenario and this could also be seen in the other parking rate scenarios. Moreover, the total idle time is the highest among all parking rates.

With regards to 0.05 euro/time step for parking price, two scenarios have the same value of profit but the optimal solutions are different which could be seen from the column “Paid parking time” and the “Total travel distance”. The static travel time scenario pays 0.3 euro more for 6 time-step parking and the dynamic one pays 0.3 € more because of the extra 3 km driving distance. This also indicates that in the dynamic scenario the model leads to a higher driving distance to avoid congestion and also avoid paying parking time in order to get a higher system profit.

When the parking price is 0.1 euro/time step, the system decides to serve 15 out of 20 requests for both of the scenarios. But the values of profit are different because of the parking cost. When ATs park at nodes except the central node 5, the system should pay for it. As a result, the dynamic travel time scenario paid 0.1€ more. In addition to this, there are 2% congestion links but no congestion delay for the passengers with dynamic travel time. This is due to the objective function which proposes the penalty for passengers’ delay but no extra costs when empty ATs are travelling slower because of the congestion.

3.5 Conclusion

This chapter proposed a mathematical model to study a dynamic travel time based AT system to provide transport service within the city area. The contribution of this chapter is to consider the travel time on each road link according to the number of vehicles travelling on and do system optimization with the objective to maximize the total profit of such system by deciding on each AT's routing selection. This model is able to include traffic congestion when computing the vehicle routing to the ATs to achieve system optimum in the vehicle routing solutions. The model was then applied to a case study with 10 time-step service period and 20 travel requests generating from the road network with 9 nodes and 12 links. From that application we were able to take the following conclusions:

Traffic congestion happens when vehicles travelling on the network and this leads to arrival delay for the trips compared with the scenario when considering the travel times static and never change on that link. The system profit is sensitive to the parking price and the penalty cost for arrival late because traffic congestion may generate benefit loss and no congestion leads to more parking costs. Sometimes the system would like to take a longer route and drive more distance in order to avoid traffic congestion and parking costs according to the profit maximization objective.

The next step will be applying the model in a real-size case with a rolling-horizon method which intends to divide one day into several horizons and make it possible to solve the real-time demand.

Chapter 4

Optimizing the dial-a-ride problem of ATs with real time demand

In Chapter 2 and 3, the demand is needed to be pre-known when solving the DARP of the AT system. However, the automated taxis should be able to serve not only the pre-booked demand but also the real-time demand. In this chapter, a rolling-horizon framework is integrated with automated taxis' dial-a-ride problem, to divide one day into several horizons in which both the pre-booked demand and the real-time demand will be considered together. Moreover, it helps reduce the problem size for the optimization problem of matching the passengers to the vehicles. With the rolling-horizon framework, it is possible to apply the model to a real-size case study city of Delft, The Netherlands.

The chapter is structured as follows: Section 4.1 reviews the existing research about dynamic DARP problem. Section 4.2 introduces the rolling-horizon framework for solving the dynamic DARP problem. Section 4.3 and section 4.4 apply the models to the case study of the city of Delft, The Netherlands and show the optimal results from the case study. Finally, main conclusions drawn from the model application are presented in section 4.5.

This chapter is an edited version of the following paper:

Liang, X., Correia, G.H. de A., van Arem, B., 2018. Applying a Model for Trip Assignment and Dynamic Routing of Automated Taxis with Congestion: System Performance in the City of Delft, The Netherlands. Transportation Research Record: Journal of the Transportation Research Board 036119811875804. doi:10.1177/0361198118758048

4.1 Introduction

The model we propose in the previous chapter is defined as a DARP of ATs, which involves transporting people from their origins to their destinations within requested time windows. It is a variation from the classic VRP, which is to design the best routes to provide services from a depot to some customers distributed in the network. In real-world applications, an important dimension of VRP is the availability of information (Psaraftis, 1988). If the assumed inputs to the VRP do not change during the solving period, or during the routing results period of implementation, this routing problem is defined as a static VRP. On the contrary, a dynamic VRP deals with a problem in which the "inputs may (and, generally, will) change (or be updated) during the implementation of the algorithm and the eventual implementation of the route" (Psaraftis, 1980). This type of VRP is also referred to as an online or real-time problem. Static DARPs assume that all passengers' requests are pre-known to the implementation of the algorithm that solves it. In a dynamic DARP case, a new customer request is eligible for consideration at the time it appears, even if it is later than the start of the vehicles' operation (Cordeau and Laporte, 2007). This matches what is required from the service that we are studying. Consequently, we will focus on a dynamic DARP problem.

In a dynamic DARP problem, vehicle routes are redefined from time to time, which requires technological support for the real-time information exchange between the vehicles and the operation centre, e.g. the position and the occupancy of the vehicles (Pillac et al., 2013). A human-driven taxi may be more difficult to route since the decisions are made involving the vehicle status information, the taxi driver and the operation centre. In an AT case, the process is simplified due to the absence of the human drivers and their corresponding stochastic drivers' decisions.

However, more attention has been paid to the algorithms for solving this kind of problems in the last two decades due to the evolution of computer power and the proposal of several heuristic and meta-heuristic methods. Given the instantaneity of the system input and the uncertainty of the vehicle routes, dynamic DARP is a more complex problem to solve. In general, methods to transfer a static VRP to a dynamic one can be divided into two categories: periodic re-optimization and continuous re-optimization (Pillac et al., 2013).

Periodic re-optimization is to return to the solving procedure each time the demand is updated. This update can be an event occurrence (a new request appears, or another one leaves), which is defined as event trigger; while it can also be a pre-determined duration, of which the most commonly used one is called the rolling horizon. Each time the optimal results are computed again with no connection to the ones in the previous computation. Yang et al. (1999) presented a rolling horizon framework for the truck fleet assignment and scheduling problem when dealing with real-time information. Luo and Schonfeld (2011) compared the rolling horizon strategy with immediately inserting requests in the dial-a-ride problem. They concluded, from their computation results, that when satisfying all the demand, the rolling horizon strategy reduced up to 10% the number of vehicles when compared with the

immediate strategy. This is because the rolling horizon strategy benefits from having information available in advance. Since the planning horizon is divided into multiple small periods of time, it is possible, in some cases, to use exact methods. This is also a way to handle the NP-hardness of the problem while abdicating from finding a global optimum for the whole period of optimization.

Continuous re-optimization performs the optimization with the available inputs and maintains information on good solutions in memory. It involves running the static VRP for the whole planning period to initialize the process and update the current routing depending on local operations, e.g. insertion heuristics, in which a new customer is inserted within the current routing, without changing the visit sequence already planned (Luo and Schonfeld, 2011). As a consequence, this process matches the nature of heuristics approaches. Dynamic VRPs are frequently solved by means of heuristics and meta-heuristics and the reason is that they can quickly compute a solution to the current problem even though they only guarantee a local optimum instead of a global one.

The periodic re-optimization solves the dynamic VRP by decomposing the problem's time dimension and generating a series of static VRP, while the continuous re-optimization focuses on the initial optimum, trying to update the solution for the local optimal routing of the real-time demand. The improvement of the heuristic or meta-heuristic algorithms to deal with this class of dynamic VRP is not the purpose of this research. We focus on formulating a mathematical model to handle the dynamic DARP for AT systems with traffic congestion and propose a rolling horizon framework to consider the real-time demand.

4.2 Rolling horizon framework

The mathematical model described in the previous section entails knowing all the demand before the optimization. This means all taxi system users should reserve this transport service and provide all the travel information including the OD and the departure time beforehand. Therefore, it would be impossible to change the ATs' movement and serve real-time demand. Using a rolling horizon framework it is possible to consider those requests and at the same time reduce the complexity of the model that was proposed above. The new notations used in this section are presented in Table 4.1.

4.2.1 Rolling horizon Framework setting

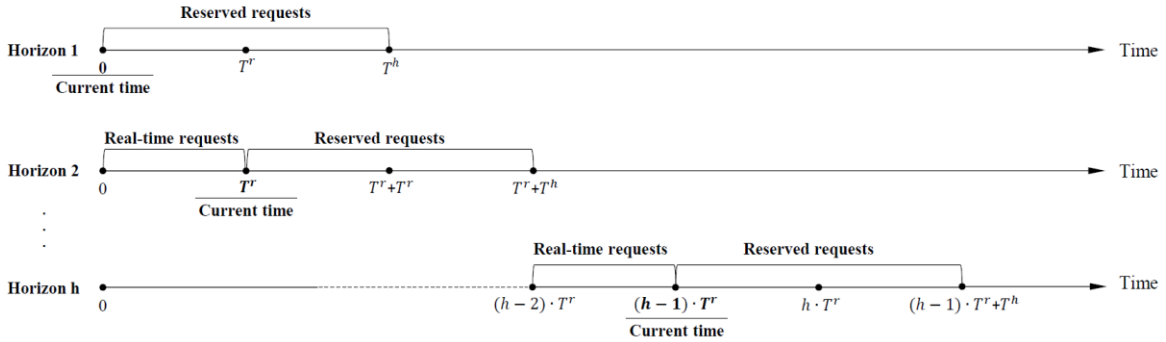
We introduce a rolling horizon framework to divide the whole optimization period into several horizons, where both the real-time and the pre-booked demand are considered. This framework uses the model [$\boldsymbol{LIP}$]: (3.1)-(3.40) in every horizon with updated demand from the previous horizon to obtain the AT routes. After that, the optimization horizon rolls forward with a specific rolling length and reaches the next horizon with updated AT requests.

Table 4.1 Notations

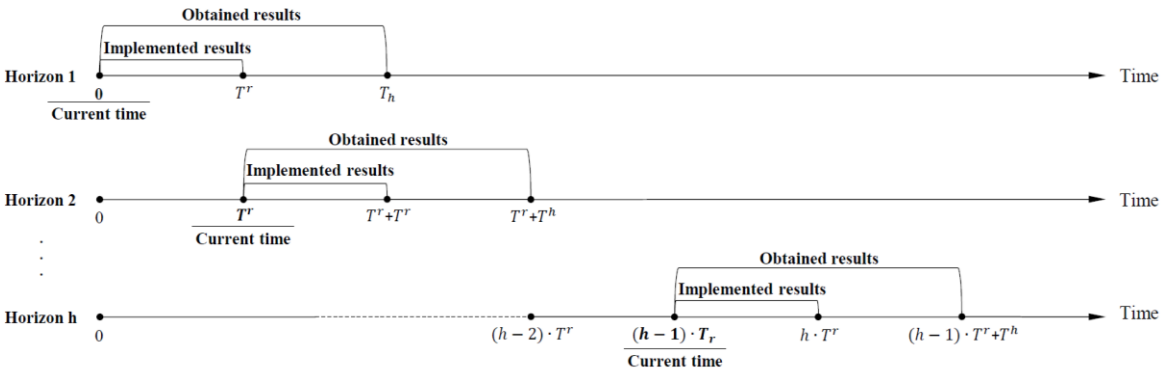
Notation	Description
Sets	
$\boldsymbol{H}$	$= \{1, \dots, h, \dots H\}$, set of horizons in an operation day, where H is the total number of horizons
$\boldsymbol{E'}$	set of the requests which are partially implemented from the previous horizon
$\boldsymbol{E^h}$	set of the requests which belong to horizon h, $\forall h \in \boldsymbol{H}$
$\boldsymbol{T^h}$	$= \{0, \dots, t, \dots, T^h\}$, set of time instants in a horizon, where T^h replaces T in the model [$\boldsymbol{LIP}$]: (3.1)-(3.40)
Parameters	
T^r	time length for rolling, $T^r < T^h$
c_{p1}	penalty cost per trip if a reserved trip is rejected by AT and needs public transport, euros/request
c_{p2}	penalty cost per trip if a real-time trip is rejected by AT and needs public transport, euros/request
loc^v	location of vehicle v at or closest to the end of the implemented period, $\forall v \in \boldsymbol{V}$
int^v	time instant when vehicle v is available in next horizon, $\forall v \in \boldsymbol{V}$
veh^e	the vehicle that satisfies request e in the current horizon, $\forall e \in \boldsymbol{E^h}, \forall h \in \boldsymbol{H}$
$st_{i_t}^v$	equals to 1 if vehicle v is travelling on a link which will finish in next horizon from time instant t to $t+1$, otherwise 0, $\forall v \in \boldsymbol{V}, \forall t \in \boldsymbol{T^h}, t < T_h, i = loc^v$
$sg_{i_t}^v$	equals to 1 if vehicle v is travelling on a link which will finish in next horizon from time instant t to $t+1$ and will end this trip at time instant $t+1$, otherwise 0, $\forall v \in \boldsymbol{V}, \forall t \in \boldsymbol{T^h}, t < T^h, i = loc^v$
$\tilde{x}_{i_{t_1}, j_{t_2}}^v$	value of the variable $x_{i_{t_1}, j_{t_2}}^v$, $\forall (i_{t_1}, j_{t_2}) \in \boldsymbol{M}, \forall v \in \boldsymbol{V}$
$\tilde{P}^{evt}$	value of the variable P^{evt}, $\forall e \in \boldsymbol{E^h}, \forall v \in \boldsymbol{V}, \forall t \in \boldsymbol{T^h}, \forall h \in \boldsymbol{H}$
$\tilde{A}^{evt}$	value of the variable A^{evt}, $\forall e \in \boldsymbol{E^h}, \forall v \in \boldsymbol{V}, \forall t \in \boldsymbol{T^h}, \forall h \in \boldsymbol{H}$
μ	expansion coefficient, representing the number of taxis with the same characteristics.

Figure 4.1 shows the rolling horizon framework. Time instant 0 is defined as the time where the ATs start to move. For horizon 1, the system collects the reserved requests which have the desired departure time within the time horizon 0 to T^h as reserved requests, and optimizes the ATs movements using these requests. After obtaining the optimal results, the AT system only implement the routing results for each vehicle from 0 to T^r and rolls forward to the next horizon. For the second horizon, two types of demand are both entered in the system. In addition to the reserved requests from T^r to $T^r + T^h$, the AT company also receives real-time demand to be served from 0 to T^r as real-time requests (these cannot be served in the previous implementing period in which they were generated due to the vehicles' routes already being optimized). The optimization process for horizon 2 will be done just before the time instant T^r, and all the ATs only perform the optimal results for $T^r$to $T^r + T^h$. Similarly,

when it rolls to horizon h, the system input will be the reserved requests for $(h-1)\cdot T^r$ to $(h-1)\cdot T^r + T^h$ and the real-time requests for $(h-2)\cdot T^r$ to $(h-1)\cdot T^r$. Then the system will only implement the routing results for $(h-1)\cdot T^r$ to $h\cdot t_r$ out of all the potential results for $(h-1)\cdot T^r$ to $(h-1)\cdot T^r + T^h$ which is the obtained results of horizon h.



(a): the input



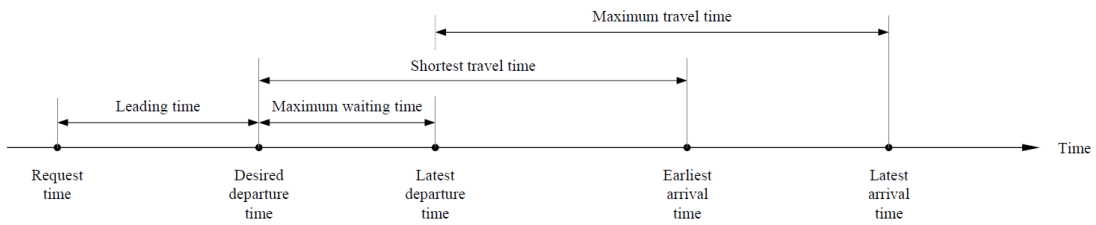
(b): the output

Figure 4.1 Rolling horizon framework

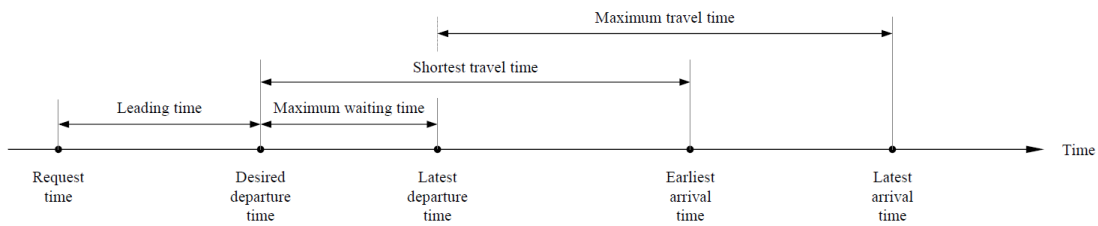
The rolling length and the horizon length can be explained from a practical respective. This rolling horizon framework is able to optimize the ATs' routing time after time with real-time information. The horizon length determines how far the system will include the pre-known information: the reserved taxi service requests. The rolling length indicates how often the system will input the new requests, update the optimal results and implement the routing plan. The rolling length is set shorter than the horizon length because the system can plan for the future requests and update once new requests enter. If the horizon is as short as a rolling length, the system cannot relocate for the next horizon trip which may decrease the system efficiency. If the rolling is as long as the horizon, the real-time request clients may wait for a whole horizon to be calculated for optimization.

4.2.2 Demand initialization

According to Figure 3.2, different time components for reserved and real-time requests can be seen in Figure 4.2. The reserved requests arrive in advance before the horizon start time. No waiting time is allowed in order to provide a better service for these clients. This means the system should serve them as they require. As a result, the desired departure time is also the earliest departure time for this kind of demand. For the real-time requests, the time they make requests is the desired departure time. Therefore, we allow some waiting time for each request to be served by ATs, since these requests are considered later than they desired. The earliest departure time is the start time of the current horizon. In this model, the maximum waiting time is set equal to the time length of rolling.



(a) reserved requests



(b) real-time requests

Figure 4.2 Time components for each request

Firstly, the desired departure time for the requests which will be analysed in horizon h should be transformed from absolute time to relative time in that specific horizon by equation (4.1).

$$(a^e)^h = a^e - (h-1) * T^r, \quad \forall e \in \boldsymbol{E^h}, \forall h \in \boldsymbol{H} \tag{4.1}$$

The departure and arrival time window for reserved requests can be calculated by equations (4.2)-(4.3).

$$(a^e)^h \leq t \leq (a^e)^h + (w^e)^h, \quad \forall e \in \boldsymbol{E^h}, \forall h \in \boldsymbol{H} \tag{4.2}$$

$$(a^e)^h + opt^e \leq t \leq (b^e)^h, \forall e \in \boldsymbol{E^h}, \forall h \in \boldsymbol{H} \tag{4.3}$$

The departure and arrival time window for real-time requests can be calculated by equations (4.4)-(4.5).

$$0 \leq t \leq (a^e)^h + (w^e)^h, \quad \forall e \in \boldsymbol{E^h}, \forall h \in \boldsymbol{H} \tag{4.4}$$

$$0 + opt^e \leq t \leq (b^e)^h, \forall e \in \boldsymbol{E^h}, \forall h \in \boldsymbol{H} \tag{4.5}$$

4.2.3 Continuity

There are three types of satisfied requests in each optimization horizon, which have different ways to be handled in the rolling horizon framework (Figure 4.3). If the optimal departure time and the arrival time for each request are both within the implementing period $(0 - T^r)$, the system will fully implement the routing results for that request (type 1). If the departure and arrival time are both after the end of the implement period T^r, the system will leave it in the demand pool and do the optimization again for it in the next horizon (type 2). If the departure time is before T^r and the arrival time is after (type 3), then the system will implement the routing results from the departure time until T^r. Based on the optimal solution, the system will know which vehicle satisfies this request and the location of it at time instant T^r. This location will be the new origin of that request and this time is the new departure time for it in the next horizon. In some cases, at time instant T^r, the vehicle v is in the middle of a road link. The cut-off time for this request will be the one closest to T^rafter T^r, and the new origin of it will be the location of vehicle v at that time instant.

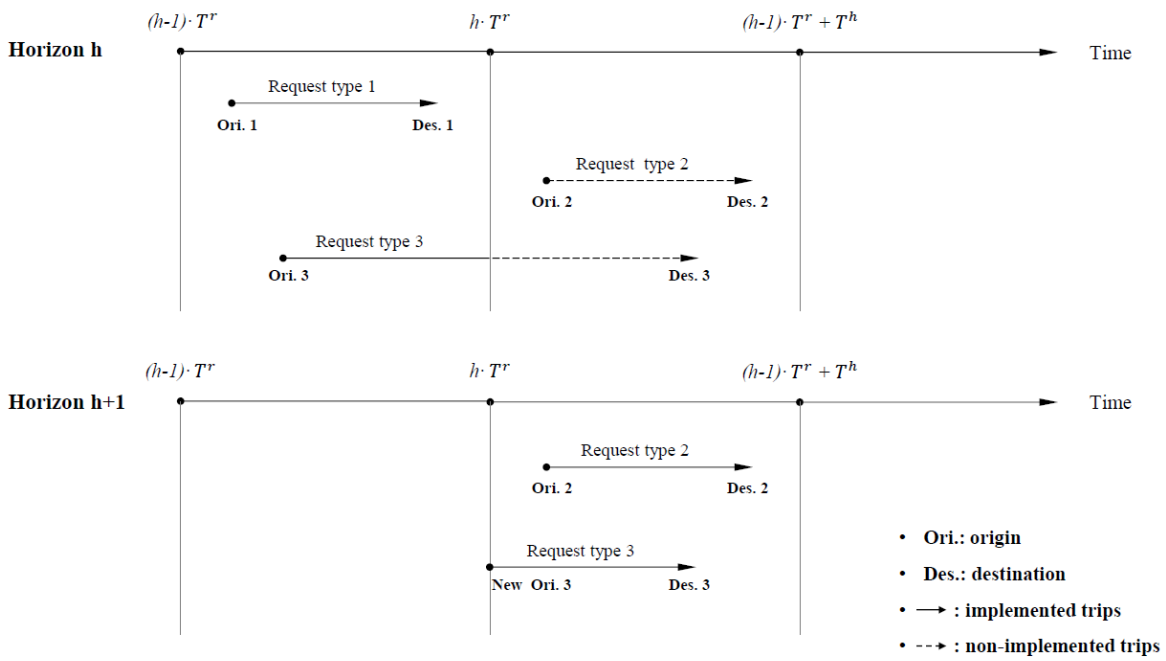


Figure 4.3 Implementing results

If the proposed model [$\boldsymbol{LIP}$] in chapter 3 is applied to the first horizon then we are able to calculate the following parameters' values, which are crucial in guaranteeing model's continuity.

$$loc^v = \sum_{\substack{(i_{t_1}, j_{t_2}) \in \boldsymbol{M} \\ t_1 < T^r,\, t_2 \geq T^r}} \tilde{x}^v_{i_{t_1}, j_{t_2}} \cdot j \qquad \forall v \in \boldsymbol{V} \tag{4.6}$$

$$int^v = \sum_{\substack{(i_{t_1}, j_{t_2}) \in \boldsymbol{M} \\ t_1 < T_r,\, t_2 \geq T^r}} \tilde{x}^v_{i_{t_1}, j_{t_2}} \cdot (t_2 - T^r) \qquad \forall v \in \boldsymbol{V} \tag{4.7}$$

$$veh^e = \sum_{v \in \boldsymbol{V}, t \in \boldsymbol{T^h}} \tilde{P}^{evt} \cdot v \qquad \forall e \in \boldsymbol{E^h} \tag{4.8}$$

Equations (4.6) calculate the final location of vehicle v at the end of the implementation period T^r. If this vehicle is in the middle of a link, then the location will be the destination of the implemented period. Equations (4.7) equal to 0 if vehicle v ends its trip and becomes available at the end of the current horizon. If not, it will be the first available time instant when vehicle v ends travelling its link in the next horizon. Equations (4.8) indicate for travel request e which vehicle satisfies it. The values of $st^v_{i_t}$ are obtained as follows: if the first available time of vehicle v is later than the beginning of the next horizon ($int^v > 0$), then the value of $st^v_{i_t}$ equals to 1, for $i = loc^v$; otherwise 0. If t to $t+1$ is the last time step vehicle v is finishing travelling on a link and it will be released at $t+1$, then the value of $sg^v_{i_t}$ equals to 1; otherwise 0.

4.2.4 The updated model

Based on the model and the system status of the previous horizon given by (4.6)-(4.8), the updated model implemented in horizon h under the rolling horizon framework is defined as follows.

$$[\boldsymbol{LIP^h}] \quad \max Z^h$$
$$= Pr \cdot \sum_{\substack{e\in \boldsymbol{E}, v\in \boldsymbol{V} \\ i=ori^e \\ j=des^e}} S_{ij}^{ev} \cdot opt_{ij} - c_f \cdot \sum_{\substack{(i_{t_1}, j_{t_2})\in \boldsymbol{A} \\ v\in \boldsymbol{V}}} x_{i_{t_1}, j_{t_2}}^{v} \cdot d_{ij} - c_v \cdot V - c_q$$
$$\cdot \sum_{\substack{i\in \boldsymbol{N}, v\in \boldsymbol{V} \\ t\in \boldsymbol{T^h}, t<T^h}} y_{i_t}^{v} - c_p \cdot \sum_{e\epsilon \boldsymbol{E^h}} \left(1 - \sum_{\substack{v\in \boldsymbol{V} \\ i=ori^e \\ j=des^e}} S_{ij}^{ev}\right) - C_w \cdot \sum_{\substack{e\epsilon \boldsymbol{E^h}, v\in \boldsymbol{V} \\ i=ori^e \\ j=des^e}} \phi_{ij}^{ev} - c_d \tag{4.9}$$
$$\cdot \sum_{\substack{e\epsilon \boldsymbol{E^h}, v\in \boldsymbol{V} \\ i=ori^e \\ j=des^e}} \left(\sum_{t\in \boldsymbol{T^h}} \left(A_{ij}^{evt} \cdot t\right) - \sum_{t\epsilon \boldsymbol{T^h}} \left(P_{ij}^{evt} \cdot t\right) - opt_{ij} \cdot S_{ij}^{ev}\right)$$

Subject to: (3.2)-(3.12), (3.16)-(3.35) plus

$$\sum_{\substack{(i_{t_1}, j_{t_2})\in \boldsymbol{M} \\ t_1\le t,\, t_2>t}} x_{i_{t_1}, j_{t_2}}^{v} + \sum_{i\in \boldsymbol{I}} y_{i_t}^{v} + \sum_{i\in \boldsymbol{I}} st_{i_t}^{v} = 1 \qquad \forall v \in \boldsymbol{V}, \forall t \in \boldsymbol{T^h} \tag{4.10}$$

Constraints (4.10) are modified from constraints (3.13), indicating that one AT can only have one status out of two: driving on a link from the current horizon, or driving on a link which has not been finished in the previous horizon.

$$\sum_{l_{t_2}\epsilon\{l_{t_2} | (l_{t_2}, i_{t_1})\epsilon \boldsymbol{M}\}} x_{l_{t_2}, i_{t_1}}^{v} + y_{i_{t_1-1}}^{v} + sg_{i_{t_1-1}}^{v} = \sum_{j_{t_3}\epsilon\{j_{t_3} | (i_{t_1}, j_{t_3})\epsilon \boldsymbol{M}\}} x_{i_{t_1}, j_{t_3}}^{v} + y_{i_{t_1}}^{v} \qquad \forall i \in \boldsymbol{I}, \forall t_1 \in \boldsymbol{T^h}, \forall v \in \boldsymbol{V} \tag{4.11}$$

Constraints (4.11) are an update of constraints (3.14). Vehicles arriving at i_t are not only from the trips in the current horizon, but also from the previous one.

$$\sum_{j_{t_2}\epsilon\{j_{t_2} | (i_{t_1}, j_{t_2})\epsilon \boldsymbol{M}\}} x_{i_{t_1}, j_{t_2}}^{v} + y_{i_{t_1}}^{v} = 1 \qquad \forall v \in \boldsymbol{V}, i = (loc^v)^h, t_1 = (int^v)^h \tag{4.12}$$

Constraints (4.12) impose that all vehicles must start from the same location in which they have stayed in the previous horizon, which replace constraints (3.15).

$$F_{ij}^{t_1} = \left(\sum_{\substack{t_1, t_2\in \boldsymbol{T_h} \\ v\in \boldsymbol{V}}} x_{i_{t_1}, j_{t_2}}^{v}\right) \cdot \mu \qquad \forall (i_{t_1}, j_{t_2}) \in \boldsymbol{M} \tag{4.13}$$

$$F_{ij}^{t} = \sum_{k\in \boldsymbol{K}} \lambda_{ij}^{t,k} \cdot k \cdot \mu \qquad \forall (i,j) \in \boldsymbol{G}, \forall t \in \boldsymbol{T}, t < T \tag{4.14}$$

Constraints (4.13) and (4.14) compute the link flow as the total number of ATs travelling on link (i, j) within one horizon, which is an update of constraints (3.16) and (3.18). Sampling expansion coefficient is used to let each AT represent μ real ATs following the same concept proposed by (Correia and van Arem, 2016a).

$$P_{ij}^{evt} = 1 \qquad \forall e \in \boldsymbol{E'}, v = (veh^e)^h, t = (int^v)^h, i = ori^{eh}, j = des^e \tag{4.15}$$

Constraints (4.15) guarantee that the partially-implemented requests from the previous horizon must be served continuously.

The following pseudo-code shows the solving process under the rolling horizon framework.

Step 0: Initialize the locations of ATs
set $Label(e) = 0, \forall e \in \boldsymbol{E^h}, h = 1$
Step 1: Filter the demand for horizon h
$\forall e \in \boldsymbol{E^h}, Label(e) = 0$
If $(h-2) \cdot T^r \leq (a^e)^h < (h-1) \cdot T^r$, then $Label(e) = 1$
end-if
Step 2: Run model [$\boldsymbol{LIP^h}$] with objective function (4.9),
subject to (3.2)-(3.12), (3.16)-(3.35), and (4.10)-(4.13), $\forall e \in \boldsymbol{E^h}, Label(e) > 0$ and (4.15), $\forall e \in \boldsymbol{E^h}, Label(e) = 2$
Step 3: Save the vehicle routing information according to (4.6) and (4.7)
Step 4: Save the request satisfying information
$\forall e \in \boldsymbol{E^h}, Label(e) > 0$
Step 4.0: If $\sum_{t,v}(\tilde{P}^{evt} \cdot t) < T^r$ and $\sum_{t,v}(\tilde{A}^{evt} \cdot t) \leq T^r$, then go to Step 4.1
else if $\sum_{t,v}(\tilde{P}^{evt} \cdot t) < T^r$ and $\sum_{t,v}(\tilde{A}^{evt} \cdot t) > T^r$, then go to Step 4.2
else if $\sum_{t,v}(\tilde{P}^{evt} \cdot t) \geq T^r$, then go to Step 4.3.
end-if
Step 4.1: Save the values of $\sum_{t,v}(\tilde{P}^{evt} \cdot t)$ and $\sum_{t,v}(\tilde{A}^{evt} \cdot t)$ as the final departure and arrival time of request e
set $Label(e) = -1$
$e = e + 1$, go to Step 4.0
Step 4.2: Save the continuity information of each partially implemented request e
Step 4.2.1: Save the AT's number who serves this request $\tilde{v} = veh^e$ according to (4.8)
set the new origin of this request $(m^e)^{h+1} = loc^{\tilde{v}}$
set the new time schedule of this request $(a^e)^{h+1} = int^{\tilde{v}}$, $(w^e)^{h+1} = 0$
Step 4.2.2: Set $Label(e) = 2$
$e = e + 1$, go to Step 4.0
Step 4.3: Set $Label(e) = 0$
$e = e + 1$, go to Step 4.0
Step 5: If $h = H$ then finish
otherwise, $h = h + 1$, go to Step 1

4.3 Case study

Several experiments were done with the ATs' DARP problem considering the impact of traffic congestion under the rolling horizon framework. In this section, we present a small case study and an application includes a large number of variables and constraints, which is the most important challenge in this research.

4.3.1 Small-scale example

We use the same example in section 3.3 to illustrate the rolling horizon framework work for the reserved requests and real-time requests. The road network and all the cost parameters remain the same. We set two horizons for this small-scale example, with 12 time steps for one horizon and 6 time steps for the rolling length. In section 3.3, there are 20 requests desiring to depart from time instant 0 to 12. We add 10 reserved requests for time period 12 to 18 and 10 real-time requests for time period 0 to 6, using the same Monte Carlo simulation in section 3.3.

The graphical results of each AT's movement for two horizons are shown in Figure 4.4. In the first horizon, the system only considers the reserved requests desiring to depart from time instant 0 to 12. The model implements the optimal results in the implemented time period, with doing nothing for the non-implemented period. To be specific, vehicle 1 will satisfy the trip from node-time 9_2 to 5_4 (request type 1 in Figure 4.3) in horizon 1. But the movements after time instant 6 including serving the request $(5_7, 4_9)$ will not be guaranteed in the next horizon. In horizon 1, vehicle 1 and 5 are travelling on a road link without passengers at time instant 6, where the model makes them frozen until the end of this link and become available for computation at time instant 7. Then the optimization period rolls forward to the second horizon, which tackles the reserved requests in time period 6 to 18 and the real-time requests in 0 to 6. Vehicle 4 serves a request from time instant 4 to 8 in horizon 1, which belongs to request type 3 Figure 4.3. In horizon 1, the system will only implement the link trips from node 1 to 5, part of its trip chain. When going to the second horizon, this request will be renewed, with a new origin of node 5, the original destination node 9 and the desired departure time 6. And constraints (4.15) guarantee that vehicle 4 is mandatory to serve this renewed request.

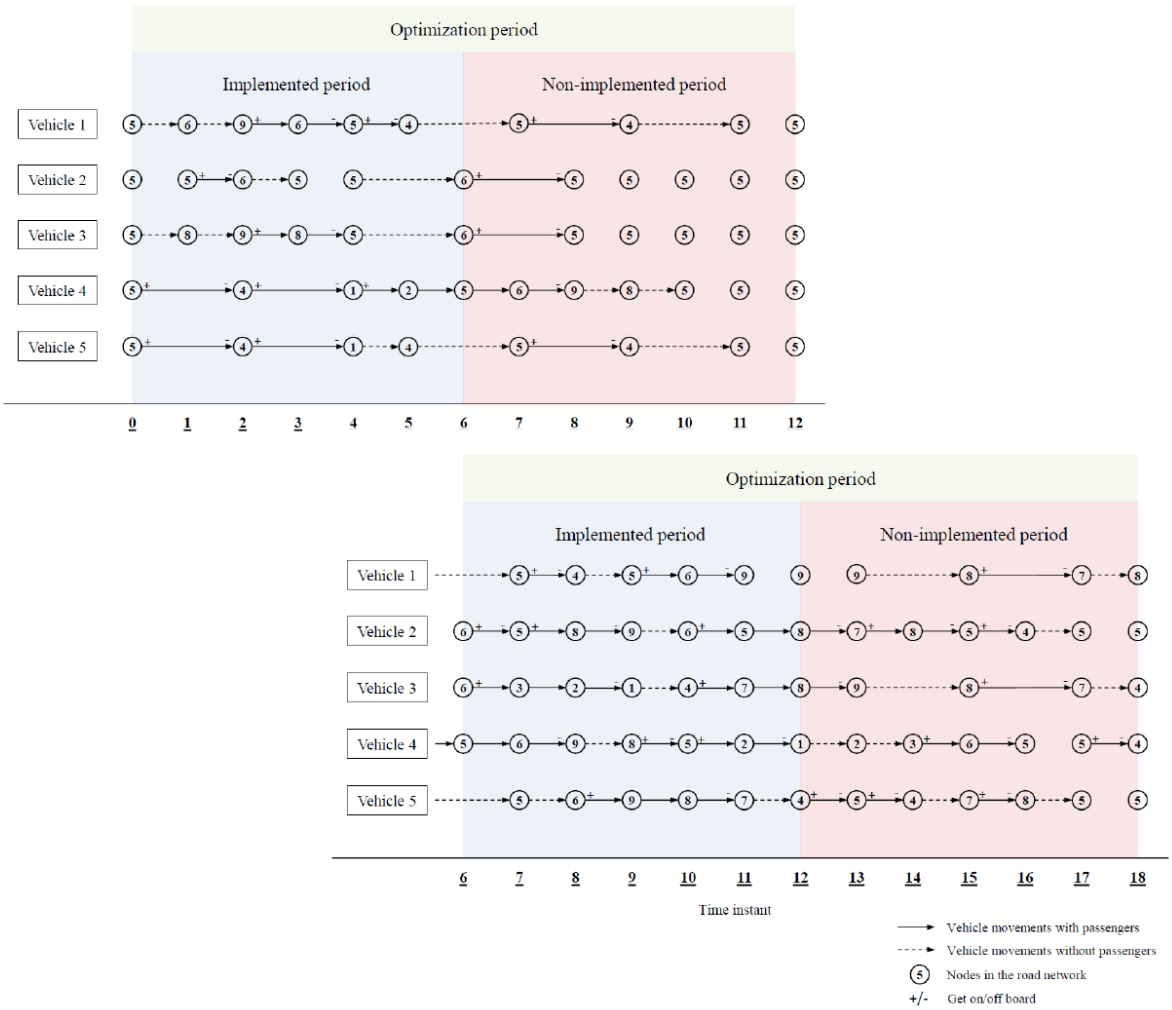


Figure 4.4 Optimization results for rolling horizon framework in small-scale example

In horizon 2, with more requests considered in the system, the vehicle movements are changed: some of the potential satisfied requests are declined; some new requests are served. Vehicle 1 satisfies request (5,4) in shorter travel time. Vehicle 2 and 3 serve more requests than in horizon 1 in the time period 6 to 12. Vehicle 4 firstly satisfies the same requests from the previous horizon, then provides services to the new requests. Vehicle 5 re-schedules its route and decides not to serve the request between $(5_7, 4_9)$ but $(6_8, 7_{11})$. The optimal results in time period 12 to 18 are also going to be updated in the next horizon.

4.3.2 Application set-up

The model is applied to the case-study city of Delft, which is located in the Netherlands. The municipality has a total area of 24 km2 and a population of about 100,000. Figure 4.5 shows the simplified road network of Delft based on Google Map. In order to apply the model for studying the AT system in the city of Delft, the following data is needed: (1) the information

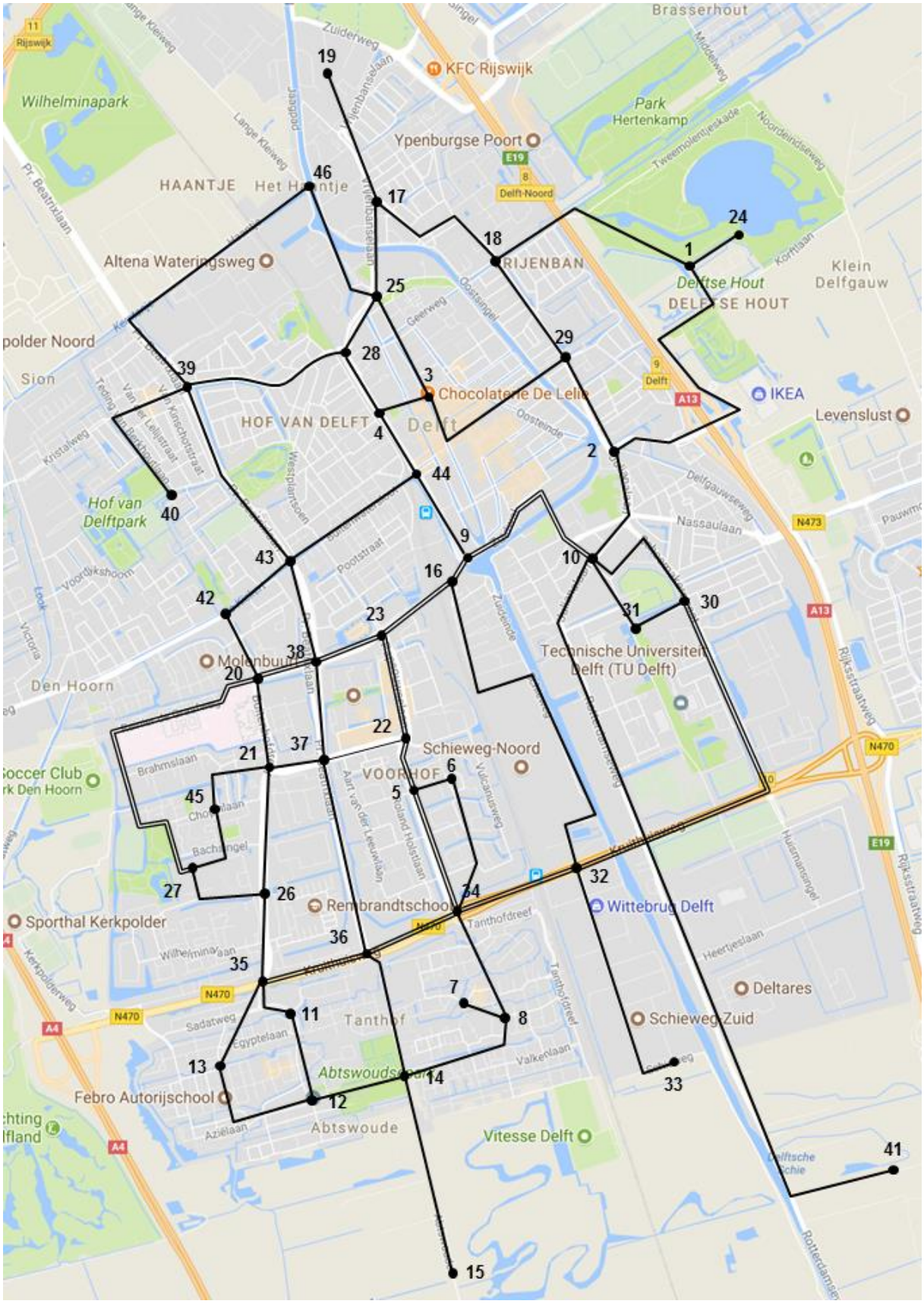


Figure 4.5 Road network of the case study

of the requests; (2) the price and cost of running the system; (3) the road network information of Delft.

The mobility data were obtained from the Dutch mobility dataset (MON 2007/2008) and transformed into the study of Correia and van Arem (Correia and van Arem, 2016a). It is provided by the Dutch government for research purpose. This data includes the origin and destination, travel mode, departure and arrival time, which are the needed information to apply the model in this chapter and it can be freely obtained online (Correia and van Arem, 2016b). This is the result of filtering the database by selecting the trips done by cars and taxis who travelled inside the city of Delft during the surveyed day.

The survey sample includes 68640 requests which were made by residents who travelled only within the city of Delft during the surveyed day. For this application, only the trips done by cars and taxis are considered. 22240 requests have resulted which are represented by 1112 requests in the operation day (expansion coefficient μ=20). At the same time, 100, 200, 300, 400 and 500 ATs available in our system represented by 5, 10, 15, 20 and 25 vehicles in the optimization model with the same value of expansion coefficient. Therefore a taxi satisfying a trip represents 20 taxis satisfying 20 requests. In these experiments, 50% of all the requests are considered to be booked in advance (reserve requests), while 50% are real-time requests. Monte Carlo simulation is used to decide which requests are pre-booked and which are real-time. The original origin and destination, as well as the desired departure time recorded in the data base, are used as preferred in the experiments.

According to the time distribution of all the requests, the earliest trip started at 6:48 and the latest trip at 23:35. Then 6:30-24:00 is chosen as the operation period of the AT service. The time step of the optimization is 2.5 min. Each horizon contains 12 time steps (T^h=12) and the rolling length is 6 time steps (T^r=6). In practice, this means each time the system looks forward to the requests in the next 30 min and updates the ATs' status every 15 min. This also means the real-time requests will wait 15 min as the maximum waiting time.

The cost values considered are as follows:

- c_r=1 euro/min: this is referred to the price rate of Uber in Amsterdam and Rotterdam in the Netherlands ("Uber," 2017). The fare is 1€ as base fare plus 1.10 €/km plus 0.25 €/min. We calculate the fare based on the shortest travel time, which means only travel time dimension of price rate is sufficient.
- c_f=0.1 euro/km: this is referred to Correia and van Arem (Correia and van Arem, 2016a)
- c_v=17.5 euro/vehicle/day: this is referred to Liang et al (Liang et al., 2016)
- c_q =0.05 euro/min: this is referred to the parking regulations in the central district of Delft, which 3 euro/hour for the underground parking garage. The parking outside the Delft city centre is free of charge according to reality.
- c_{p1}=2 euro/request: this is referred to the public transport fare in Delft.

- c_{p2}=0 euro/request: no penalty cost for the unsatisfied real-time requests. The penalty for unsatisfied reserved and real-time requests are given different values, which is mainly because the system should give priority to reservations.
- c_d=0.2 euro/min: this is referred to the price rate.
- c_w=0.04 euro/min: the penalty for the late arrival of real-time requests is relatively low since the maximum waiting time is strict: 2.5 min.

The network has 66 road links and 46 nodes Figure 4.5. Some of the links have one lane per way and some have two lanes. The capacities were considered as 1500 and 3000 vehicles per hour per direction based on the number of lanes on that link. The maximum speed was assumed to the 50 and 70 km/h respectively for the lower and higher capacity links. In Figure 4.5, the higher capacity links are shown in double lines and the lower ones are in single lines.

The shortest travel times for the requests are computed by the shortest path method with the minimum link travel time. The minimum travel time δ_{ij}^{min} on the link (i, j) is obtained from the free flow speed. The maximum travel time δ_{ij}^{max} in time step is computed based on a speed of 5 km/h, which is reflected by the value of break-point k. There are 3 break-points for the road links with lower capacity and 6 break-points with higher capacity. These travel times are obtained based on the Bureau of Public Roads function (Dafermos and Sparrow, 1969): $t = t_0(1 + a \times (\frac{V}{Q})^b)$, where t_0 is the minimum travel time, $\frac{V}{Q}$ is the value of break-point k, a is $\left(\delta_{ij}^{max}/\delta_{ij}^{min}\right) - 1$ and b is 4 for experimental purposes. The initial location of each vehicle is set to be at five nodes by constraint (3.15): node 27, node15, node 16, node 19, node 24, where they all have V/5 vehicles available to balance the geographical distribution. The shortest travel time between the nodes (only for which there is demand) is computed by the shortest path method with free flow travel time (opt_{ij}). The travel distance on each link is the link length in reality.

We test our algorithm on the Delft road network for ten scenarios. These scenarios are set with different types of travel times and various fleet sizes. Then, we will examine the results of each scenario individually.

- Scenario I to V: dynamic travel time scenarios considering traffic congestion in route designing with 100, 200, 300 400 and 500 vehicles as AT fleet in the system.
- Scenario VI to X: static travel time scenarios assuming that the travel times do not change with different levels of traffic congestion impacts and are always equal to the shortest travel time on that link. Correspondingly, The fleet size varies from 100 to 500 vehicles.

4.3.3 Computation performance

We implemented the integer programming formulation in Mosel language and solved by Xpress-Optimizer in an i5 processor @3.10GHz, 12.00 GB RMA computer under a Windows

7 64-bit operating system. Xpress is an optimization tool that uses an advanced branch-and-cut algorithm for solving integer programming problems.

There are two control parameters set in the optimizer to govern the solution procedure. Both of them are used to determine when the optimizer will terminate. The first one works when the gap between the best integer solution's objective function value and the current best bound by the global search is equal to or small than 5%. The second control parameter is to decide the maximum time the optimizer will run before it terminates is one hour if an integer solution has been found. Otherwise, it will continue until an integer solution is finally found.

We choose scenario III with 300ATs and dynamic travel time as an example to expound the computational performance. The computational time for 70 horizons in the scenario with 300 ATs varies from 38s to 3648s, showing in Figure 4.6. In average each horizon takes 19 min to find the optimal solution. The median of this set of values is 210s, which indicates that half of the horizons are calculated longer than 210s and the other half are shorter than that. 18 out of 70 horizons are reaching the maximum 1-hour computational time, meaning after 1-hour searching, the optimizer still cannot converge the gap between the best bound and the currently best integer solution to 5%. The high computation time is mainly because of the size of the formulation: a large number of variables and constraints. On one side, this is results from the mathematical formulation itself (e.g. the vehicle movement variables have five dimensions). On the other side, the number of passenger requests also has an influence on the size of the problem. Figure 4.6 shows the computational time for each horizon and the number of variables and constraints. In this model, all the variables are integer variables, which increases the complexity of the branch-and-bound. When there are more variables and constraints, the computational time tends to be longer and vice versa.

Another reason which may influence the computational time is the symmetries of the solutions, i.e. fewer symmetries of the solutions, better model. In this chapter, we formulate the system's activities by each ATs and each client, which is a non-aggregated method to reveal each vehicle's movement and serving information to the system operators. But this leads to a high level of symmetry of the feasible solutions. For example, if m requests are desired to depart from node A to node B at the same time and there are n homogeneous ATs available at node A. There is a possibility for each AT to serve any one of the requests with the same contribution to the objective function value. This can be seen in Figure 4.7, which presents the number of requests and the number of variables and constraints for every horizon. In horizon 1, there are only 6 requests but the model generates 115619 variables and 112021 constraints, being much more than most of the horizons. This is because of the symmetry of the solutions. If all the homogenous ATs start their trip from the same node, then it will take a long time to search for all the possible combinations of passengers and vehicles. We reduce the symmetries to some extent by assigning all the ATs to different initial locations, but this model still costs high computation times.

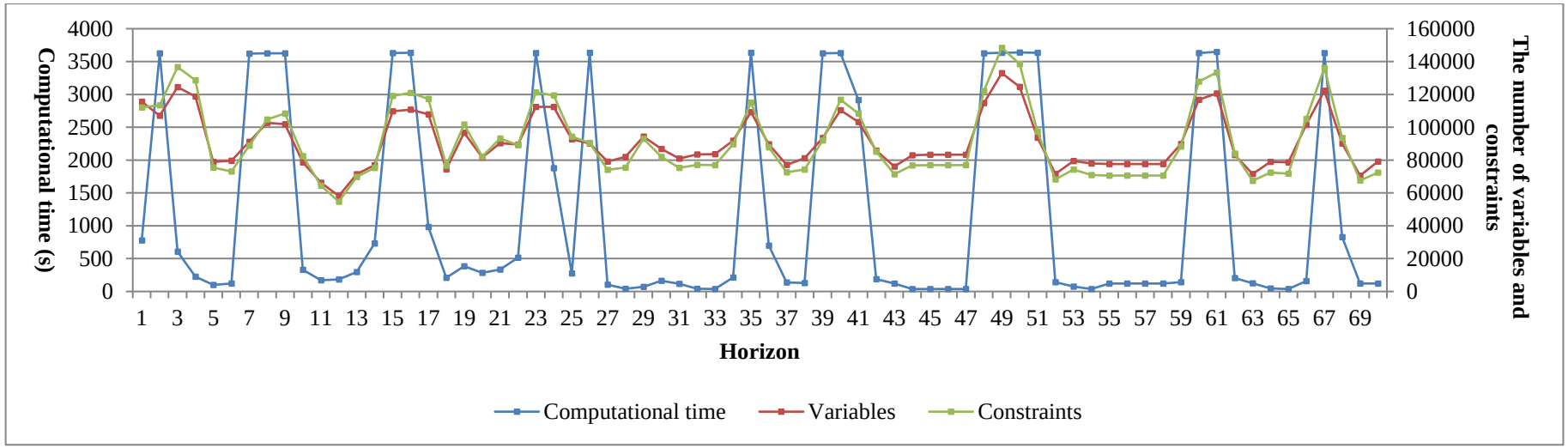


Figure 4.6 Computational time vs. the number of variables and constraints, 300 ATs, dynamic travel time

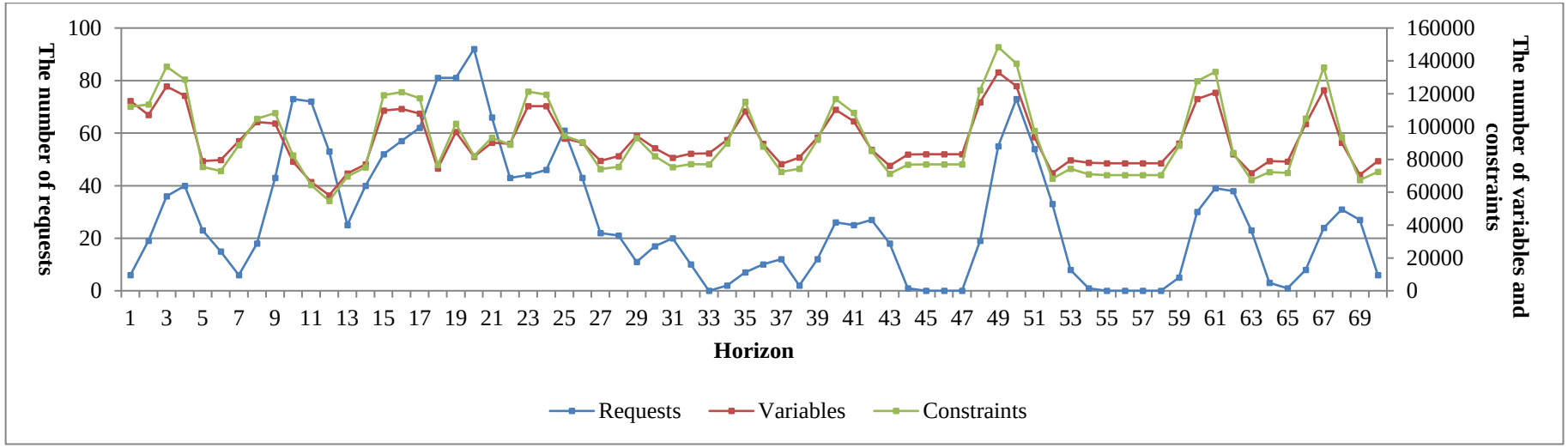


Figure 4.7 The number of requests vs. the number of variables and constraints, 300 ATs, dynamic travel time

4.4 Experiments and results

A comparison of AT routing optimization considering traffic congestion between different fleet size was done and is present in three aspects: the satisfying performance, vehicle usage, and time delay. All the travel times on each road link are considered as dynamic, meaning the traffic congestion is considered in these scenarios.

4.4.1 Fleet size variation

A comparison of AT routing optimization considering traffic congestion between different fleet size was done and is present in three aspects: the satisfying performance, vehicle usage, and time delay. All the travel times on each road link are considered as dynamic, meaning the traffic congestion is considered in these scenarios.

The values of indicators for satisfying performance for dynamic travel time scenarios with 100 to 500 ATs are presented in Table 4.2. In these five scenarios, the fleet size is a critical parameter with respect to satisfying performance. When there are only 100 vehicles in the taxi system, 1960 requests are satisfied in total. In this case, the satisfied rate of 8.8% is the lowest among all the scenarios. When the fleet size increases to 200, the satisfied demand almost doubled. This trend continues in the next two scenarios with 300 and 400 ATs, but the growth rate is falling in them. In the last scenario, the number of satisfied requests decreases, with 20 less than the 400-AT scenario. This means with 100 more ATs available in the city, the system cannot provide a better service coverage, which also means from that moment, other reasons become the restriction of the satisfaction, e.g. the road capacity and the allowed departure delay.

Table 4.2 Satisfying rate for scenario I-V

Scenario	Fleet size	Satisfied requests	Satisfied rate	Satisfied reserved requests	Satisfied rate of reserved requests	Satisfied real-time requests	Satisfied rate of real-time requests
I	100	1960	8.8%	1720	15%	240	2%
II	200	3680	16.5%	3220	29%	460	4%
III	300	4860	21.9%	3980	36%	880	8%
IV	400	5580	25.1%	4300	39%	1280	12%
V	500	5560	25.0%	4100	37%	1460	13%

The satisfying performance on different types of requests for different scenarios is consistent, with the satisfied reserved requests are always much more than the real-time requests. This is because of the different unsatisfied penalty costs. Since the penalty of reserved requests is higher than the real-time ones, the system decides to serve reserved requests as a priority. For

example, when there are 300 ATs available, 36% reserved requests are served, while 8% of real-time requests are satisfied. Therefore the value of penalty costs has a significant influence on deciding which requests to be served. It also reveals that the system can use this method to adjust the service strategy for different kinds of demand.

The vehicle usage indicates the working efficiency of each AT, with the results showing in Table 4.3. In these five scenarios, the most efficient is the one with 100 ATs, with 19.6 requests per vehicle. This also can be seen from the column “average travel time”, “average idle time”, “idle rate” and “average travel distance”. Scenario V with 500 ATs has the highest idle time and idle rate. This means more than half of the total operation time ATs are parking somewhere without any movement. The average travel distance is 147.5 km per AT, which also lower than in any other scenarios with the dynamic travel time condition. Actually, it doesn’t conflict with the unsatisfied request rate. The demand is not uniformly distributed during the day which brings out the peak horizon and the off-peak horizons. In off-peak horizons, 500 ATs are redundant and generate idle time, while in peak horizons they are not enough to serve all the requests.

Table 4.3 Vehicle usage for scenario I-V

Scenario	Fleet size	Avg. satisfied requests per vehicle	Avg. travel time min/veh.*	Avg. idle time min/ veh.	Idle rate	Avg. travel distance km/veh.	Avg. travel distance with passengers km/veh.	Travel distance with passengers/ travel distance	Avg. travel distance without passengers km/veh.	Travel distance without passengers/ travel distance
I	100	19.6	554.5	495.5	47%	260.0	165.5	64%	94.6	36%
II	200	18.4	512.8	537.3	51%	237.8	155.7	65%	82.1	35%
III	300	16.2	455.3	594.7	57%	205.0	136.6	67%	68.4	33%
IV	400	14.0	389.6	660.4	63%	171.5	115.8	68%	55.7	32%
V	500	11.1	330.3	719.7	69%	147.4	98.6	67%	48.7	33%

veh.*: vehicle

Among the average travel distance for each AT, the occupied kilometres are always more than the empty kilometres, for all five scenarios. The system only pays costs but no revenue for empty kilometres, therefore it will try to avoid travelling without passengers. With the increasing of fleet size, the travel distance with passengers is increasing slightly, until the scenario 4 with 400 ATs. For the last scenario, the percentage falls down to 67%. This indicates that with more ATs, the system has more flexibility to deploy the vehicles to serve different trips which brings the possibility to decrease the empty travel distance.

The influence of traffic congestion on AT service is quantified by time delay. The results for different kinds of time delays are presented in Table 4.4. The congestion delay is defined as

the time difference for each vehicle movement ($x^{v}_{i_{t_1},j_{t_2}}$) between the real travel time and the shortest travel time on each road link. They can be distinguished as occupied trips (with passengers) and unoccupied trips (without passengers). In the objective function, the revenue is calculated by the shortest travel time for each request. The time delay happening on the occupied trips should be charged, while the unoccupied don't need to. As a result, the congestion delay for link trips with passengers is always higher than trips without passengers. This can be illustrated by the column "Avg. congestion delay for occupied trips" and "Avg. congestion delay for unoccupied trips". It also reveals that the system can use this method to apply different charging rules to adjust the time and place the congestion happens.

Table 4.4 Time delay for scenario I-V

Scenario	Fleet size	Total congestion delay min	Avg. congestion delay min/veh.*	Total congestion delay for occupied trips min	Avg. congestion delay for occupied trips min/veh.	Total congestion delay for unoccupied trips min	Avg. congestion delay for unoccupied trips min/veh.	Total delay min	Avg. delay min/ satisfied requests	Avg. waiting time min/satisfied real-time request
I	100	500	5.0	0	0.0	500	5.0	150	0.08	11.88
II	200	1200	6.0	250	1.3	950	4.8	1900	0.52	11.85
III	300	1950	6.5	500	1.7	1450	4.8	3700	0.76	12.05
IV	400	3200	8.0	700	1.8	2500	6.3	4900	0.88	11.17
V	500	3850	7.7	950	1.9	2900	5.8	6150	1.11	11.20

veh.*: vehicle

The total congestion delay time per vehicle keeps increasing from scenario I to scenario V, but the highest average value per vehicle is 8.0 min when there are 400 ATs, and it falls down to 7.7 min. This demonstrates that in the scenario with 500 ATs, the additional 100 ATs do not generate congestion as much as the ATs in scenario Iv, meaning the vehicle usage is lower.

When there are more vehicles planning to travel on the same link, the model should decide whether to choose the shortest path for that OD pair or take a farther path to avoid traffic congestion. This is defined as detouring. For example, if a request is from node 26 to node 37 in Figure 4.5 and the shortest path is 26-21-37, the model may choose the path 26-35-36-37 with a longer travel time but avoid congestion. We compute the difference between the real travel time and shortest travel time for each request, which includes not only the congestion delay but also the extra time costs in detouring. Therefore, the column "Avg. delay for satisfied requests" is the one to reflect the impact of traffic congestion on AT service. When there are 100 ATs in the system, on average each satisfied request has 0.08 min time delay. This value keeps increasing to 1.11 min when there are 500 ATs. In average the satisfied

requests are not delayed seriously, because of the un-uniform distribution of travel demand in time and space: in-peak time horizons and high demand locations, the traffic congestion is more willing to happen than the off-peak horizon and other rural places.

4.4.2 Dynamic vs. static system

Another kind of scenarios is considered, which assumes that the travel times do not change and are thus always equal to the minimum travel time on that link, even if the traffic flow varies (static travel times). The travel time constraints and capacity constraints are turned off. Another key variable for creating these scenarios is the fleet size. The comparison results are shown in Table 4.5.

Table 4.5 Profit, satisfied rate, and time delay for scenario I to X

Scenario	Fleet size	Profit €/day	Over-estimated profit	Satisfied requests	Over-estimated satisfied requests	Avg.* delay min/ satisfied requests	Avg. congestion delay min/veh.**
I	100	10710	-	1960	-	0.08	5.0
VI		11138	4%	2000	2%	0	0
II	200	38315	-	3680	-	0.52	6.0
VII		40672	6%	3740	2%	0	0
III	300	54809	-	4860	-	0.76	6.5
VIII		69372	27%	5480	13%	0	0
IV	400	62830	-	5580	-	0.88	8.0
IX		79591	27%	6380	14%	0	0
V	500	60402	-	5560	-	1.11	7.7
X		97938	62%	7540	36%	0	0

Avg.*: average
veh.**: vehicle

Comparing the static travel time scenarios with dynamic ones, it is easy to find that the AT system is always more profitable when considering the minimum travel times. When the fleet size is 100, the system obtains 10710 € under the dynamic travel time condition, which is 4% lower than the static travel time with 11138 €. In addition, the number of satisfied trips in the dynamic travel time scenarios is also more than the one from the static travel time scenarios. With 300 ATs scenario III has 27% less profit, while with 500 ATs the dynamic travel time scenario has 62% less daily profit. This is because when the travel time is always the minimum value, the system, would take less time than the dynamic travel time scenario, which makes the ATs possible to serve more trips and receive more profit. But it cannot reflect the real traffic status and the impact of ATs on the travel time. This could be seen from

the column “Avg. delay” and “Avg. congestion delay” in Table 4.5, which represents the influence of traffic congestion on travel times, compared with no delay in the static travel time scenarios.

4.5 Conclusions

AVs have drawn great attention in the recent decades evolving from a concept model to a real one driving on the roads in some challenging studies. An emphasis has been put on technology in recent research focusing on the transport reliability and safety, and also the interactions between the AVs and others like the road network, human beings, regular vehicles and so on. Nevertheless, some questions related to the application of AVs are still existing, like their use in public transport.

This chapter proposed a mathematical model to study a dynamic travel time based AT system to provide transport service with the city area. The contribution of this chapter is to consider the travel time on each road link according to the number of vehicles travelling on and do system optimization with the objective to maximize the total profit of such system by deciding on each AT’s routing problem including real-time information. It is necessary to consider the model in a situation which involves a large number of passengers’ requests and also a large number of vehicles, which makes this model NP-hard. This model involves both the reserved requests and real-time request via rolling horizon structure and optimizes the vehicle routing horizon by horizon, which is also possible to decrease the number of variables and constraints in an optimization process and reduce the computational time to an acceptable level. The model was then applied to a case study in the city of Delft, the Netherlands with 17.5 hours service period and 23340 travel requests generating from the road network with 46 nodes and 66 links. From that application it was possible to take the following conclusions:

The fleet size is an important factor in system profitability and satisfying performance, regardless of the dynamic or static travel time. The reason is that more ATs are able to serve more requests which bring more income for the AT system, even though the depreciation costs are also increasing and the service efficiency is decreasing. When the fleet size keeps increasing, the number of satisfied requests are increasing with the same speed, which demonstrates that the fleet size is not the only reason to restrict the satisfying performance, while the road network capacity, the un-uniform distribution of the demand and the strict waiting time are all able to influence the satisfying rate.

The influence of considering travel time as dynamic in ATs’ vehicle routing problem can be seen from two aspects: the time delay on the congested road links and the extra travel time caused by detouring. Only concentrating on the congestion road links is not sufficient to analyse the impact of traffic congestion, while the system may choose to take a farther path to avoid congested links according to the objective function.

Knowing that real-life travel conditions have dynamic travel time, application of a static minimum travel time vehicle routing model will overestimate the system profitability. Traffic

congestion happens when vehicles are travelling on the network, and it leads to less profit, arrival delay and fewer satisfied requests than in the case of the static travel time. But it reflects the real traffic situation and makes the routing results more realistic.

The current version of the optimization model is not completely practice-ready, due to the computational time. But this formulation still has benefits for future work. Firstly, it can be used to develop heuristics for ATs' DARP problem, which may accelerate the optimization process. Moreover, this model can be used to do sensitive analysis such as the penalty costs to the satisfying performance and the price rate to the system profitability. Third, this model only considers individual passengers, however, it would be interesting to allow ride-sharing in the same AT for the passengers who have similar origins and destinations. Finally, we intend to involve choice modelling to analyse people's preference for using ATs as a mode of urban mobility and combine with the ATs' DARP problem.

Chapter 5

Optimizing the dial-a-ride problem of ATs with Lagrangian relaxation based solution algorithm

In Chapter 4 and 5, the proposed models considering traffic congestion are simplified by linearizing the non-linear travel time function. This sacrifices the model's accuracy to some extent. It is also time-consuming when solving this problem in a large-scale network. In addition, the automated taxi service is defined as an individual service, not considering ride-sharing. In this chapter, we keep the non-linear function of the dynamic travel times when optimizing the automated taxis' dial-a-ride problem. Moreover, this system allows passengers to share a ride with similar origins, destinations and time schedules. The model is hard to solve due to the nonlinearity and the large number of decision variables and constraints. Therefore, we propose a solution approach based on a customized Lagrangian relaxation algorithm, which enables to identify a near optimal solution for this difficult problem. The formulation and the solving algorithm are applied to the case study in Delft, The Netherlands to demonstrate the computation efficiency and the system performance.

The chapter is structured as follows: Section 5.1 reviews the existing research about using Lagrangian relaxation to solve NP-hard problems. Section 5.2 introduces the Lagrangian relaxation solution algorithm proposed in this thesis. Section 5.3 and section 5.4 apply the models to the case study of the city of Delft, The Netherlands and show the optimal results from the case study. Finally, in section 5.5, we present the main conclusions drawn from the model application.

This chapter is an edited version of the following paper:

Liang, X., Correia, G.H. de A., An, K., van Arem, B., Automated taxis' dial-a-ride problem with ride-sharing considering dynamic travel times: an application to Delft, The Netherlands (under review)

5.1 Introduction

In this chapter, a non-linear integer programming (NLIP) model and a solution algorithm are proposed to address the problem of assigning ATs to clients and define their routes on an urban road network. The model considers traffic congestion by incorporating a non-linear travel time function that varies with the flow of the ATs. The model also allows ride-sharing, in order to increase the transport efficiency of the AT system, which is different from the previous chapters. A customized Lagrangian algorithm is developed to consecutively solve the proposed NP-hard routing problem at each time-horizon for real case-study applications.

Allowing different passengers to share a ride is a new consideration in this chapter. Private or public ride-sharing aims to bring together travellers who have similar itineraries and time schedules to share rides (Agatz et al., 2012, 2011; Correia and Viegas, 2011; Schaller, 2018). The large demand and the low occupancies in private transport in peak hours create traffic congestion in many urban areas. Ride-sharing allows people to use transport capacity more efficiently (Furuhata et al., 2013). In the conventional ride-sharing system, users can provide a ride as a driver or ask for a ride as a passenger. Once the travel requests are submitted, there will be a matching between the drivers and the riders. In the matching process, the key constraint is the time schedules of the rides. The drivers should have sufficient time flexibility since they need to accomplish the pick-up and drop-off of the passengers and then arrive at their own destinations. If ATs are used in the service scheme of ride-sharing, they will provide the opportunity to transform the role of the drivers into passengers, who have no need to stay in the vehicles for the whole ride. Currently, ride-sharing is happening for example with Uber-pool systems whereby a person may request a ride at a lower price but be willing to share with other passengers.

The NLIP model can be solved using most of the commercial software like Xpress or CPLEX. But generally, the computation time is excessively long, especially when applying it to large scale problems, due to its NP-hard property. To solve a large-scale problem as described in section 3 and 4, an efficient solution algorithm is needed. In this chapter, we develop a customized Lagrangian relaxation algorithm (Fisher, 1981) to solve the proposed model (Brannlund et al., 1998; Cui et al., 2010; Mahmoudi and Zhou, 2016; Park et al., 2010; Tanatmis et al., 2010; Yang and Zhou, 2014; Zhou and Teng, 2016).

Table 5.1 Notations

Notation	Description
Sets	
$\boldsymbol{L}$	$= \{1, \dots, l, \dots L\}$, set of Lagrangian iterations, where L is the total number of iterations.
$\boldsymbol{K}$	$= \{0, \dots, k, \dots K\}$, set of traffic assignment iterations, where K is the total number of iterations.
Parameters	
$vcap$	seating capacity of the vehicle, which is the maximum number of passengers that can share a ride.
μ_1^{evt}	Lagrangian multiplier associated with constraints (5.4), $\forall e \in \boldsymbol{E^h}, \forall v \in \boldsymbol{V}, \forall t \in \boldsymbol{T^h}, \forall h \in \boldsymbol{H}$.
μ_2^{evt}	Lagrangian multiplier associated with constraints (5.5), $\forall e \in \boldsymbol{E^h}, \forall v \in \boldsymbol{V}, \forall t \in \boldsymbol{T^h}, \forall h \in \boldsymbol{H}$.
ρ^l	step size of Lagrangian relaxation, $\forall l \in \boldsymbol{L}$.
UB^l	current upper bound obtained in iteration l, $\forall l \in \boldsymbol{L}$.
LB^*	best lower bound found.
π^l	control parameter, $\forall l \in \boldsymbol{L}$.
$\tilde{F}_{ij}$	value of the variable F_{ij}, $\forall (i,j) \in \boldsymbol{G}$.
$\tilde{\delta}_{ij}$	value of travel times in time steps when travelling on link $(i\ j)$, $\forall (i,j) \in \boldsymbol{G}$.
Vol_{ij}	value of ATs' volumes when travelling on link $(i\ j)$, $\forall (i,j) \in \boldsymbol{G}$.
Decision variables	
P^{evt}	binary variable equal to 1 if travel request e is done by vehicle v starting at time instant t, otherwise 0, $\forall e \in E, \forall v \in V, \forall t \in T, a^e \leq t \leq a^e + w^e$.
A^{evt}	binary variable equal to 1 if travel request e is done by vehicle v ending at time instant t, otherwise 0, $\forall e \in E, \forall v \in V, \forall t \in T, a^e + opt^e \leq t \leq a^e + w^e + lon^e$.

Fisher (2004, 1981) showed that this approach can efficiently solve a wide range of difficult mixed integer problems, e.g. the travelling salesman problem, the scheduling problem, the location problem, the assignment problem, etc. An et al. (2017) applied it to solve a sensor location problem for object positioning and surveillance, where the Lagrangian relaxation provides a lower bound (minimization problem) to the original problem. Bai et al. (2011) introduced a Lagrangian relaxation based heuristic algorithm to solve the biofuel refinery location problem under traffic congestion and obtain the near optimal feasible solutions efficiently. Lei and Ouyang (2018) proposed a Lagrangian relaxation based algorithm to solve the one-commodity pickup and delivery problem. The numerical experiments show that it is able to generate a good solution for large-scale cases in short computation time. The pickup and delivery problem was also addressed by Imai et al. (2007) in the field of container load from/to an intermodal terminal and solved by a sub-gradient heuristic based on a Lagrangian

relaxation to identify a near optimal solution. Shen et al. (2011) studied an inventory routing problem in crude oil transportation and developed a Lagrangian relaxation approach for finding the near optimal solution of the mixed integer problem. An et al. (2017) designed a customized Lagrangian relaxation algorithm with an embedded approximation subroutine to solve the sensor location problem. All these applications demonstrate that this algorithm can be used to provide bounds in discrete optimization and can be further integrated with other methods e.g. branch-and-bound, heuristics to produce a near optimal solution to these challenging problems.

5.2 Lagrangian relaxation solution algorithm

The notations used in this section are presented in Table 5.1.

5.2.1 Integer programming models with non-linear travel time function

We update model [$\boldsymbol{LIP^h}$] proposed in chapter 4 and formulate a new model with non-linear constraints imposing travel time as a function of link flow. The changes regarding the system assumptions and settings that have been done are described as flow:

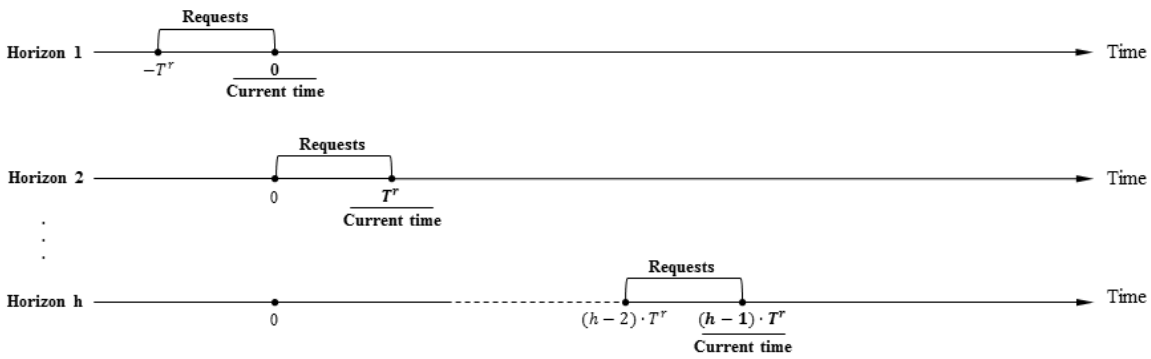
- We eliminate decision variables S_{ij}^{ev} in model [$\boldsymbol{LIP^h}$] and only use P_{ij}^{evt} and A_{ij}^{evt} to describe requests' pick-up and delivery activities. The index of the origin i and the destination j of variable P_{ij}^{evt} and A_{ij}^{evt} are also eliminated thus we indicate this information by index e.
- In chapter 3 and 4, the ATs are allowed to stay parking when they are idle, while in this chapter parking is not allowed which means that all the ATs should be cruising all the time pro-actively relocating to demand areas. The reason for making this assumption is to apply the iterative assignment solving process, which will be further explained in section 5.2.2.
- In chapter 3 and 4, the AT system only provides service to individual trips. Here the AT system allows several clients to be pooled together (ride-sharing) respecting vehicle capacity and travellers' schedules.
- In order to simplify, the delivery delay is considered to be a combination of pick-up delay (passengers' waiting time for ATs) and congestion delay (the real travel time minus the shortest travel time). This is different from chapter 3 and 4 where these two parts of cost are considered separate in the objective function. However, in this chapter, they are calculated using the same unit cost.
- In order to simplify the model, only real-time requests are considered in this chapter meaning that all the passengers would like to be picked up at the time they submit their requests. Figure 5.1 shows the rolling horizon framework in this chapter and Figure 5.2 shows the time windows for the real-time requests. Moreover, we defined

the latest arrival time as $b^e = a^e + w^e + lon^e$, which makes the departure and arrival time window for these requests as (5.1) and (5.2).

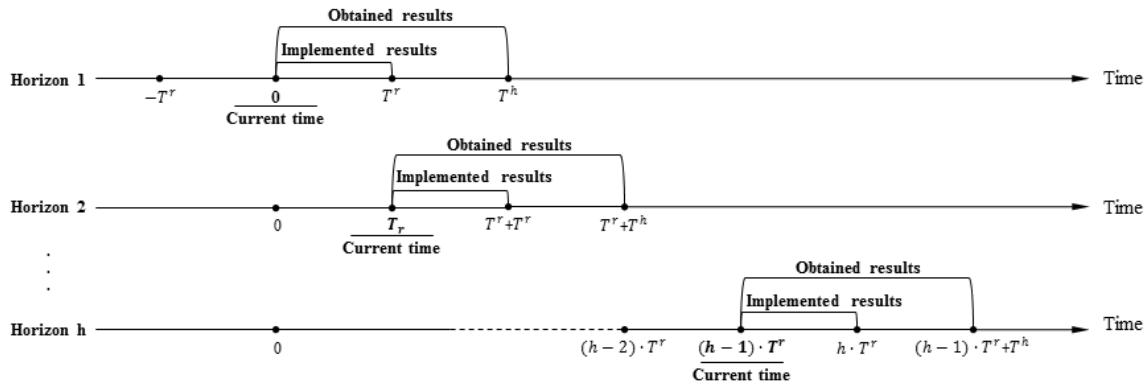
$$0 \le t \le (a^e)^h + (w^e)^h, \ \ \forall e \in \boldsymbol{E^h}, \forall h \in \boldsymbol{H} \tag{5.1}$$

$$opt^e \le t \le (a^e)^h + (w^e)^h + (lon^e)^h, \forall e \in \boldsymbol{E^h}, \forall h \in \boldsymbol{H} \tag{5.2}$$

•



(a): the input



(b): the output

Figure 5.1 Rolling horizon framework

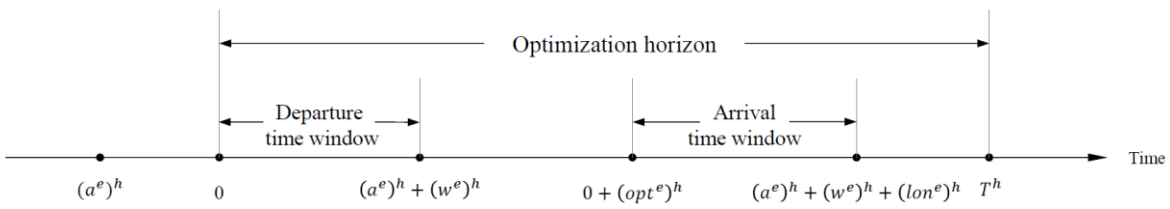


Figure 5.2 Departure and arrival time windows for the requests

The updated optimization mode $[\boldsymbol{NLIP^h}]$ with non-linear travel time constraints for solving the DARP problem has the following formulation.

$$[\boldsymbol{NLIP^h}] \quad \max Z^h = \sum_{e\epsilon \boldsymbol{E}, v\in \boldsymbol{V}, t\epsilon \boldsymbol{T^h}} c_r \cdot P^{evt} \cdot opt^e - \sum_{\substack{(i_{t_1}, j_{t_2})\in \boldsymbol{M} \\ v\in \boldsymbol{V}}} c_f \cdot x^v_{i_{t_1}, j_{t_2}} \cdot d_{ij} - c_v \cdot V - \sum_{e\epsilon \boldsymbol{E^h}} c_p \cdot \left(1 - \sum_{v\in \boldsymbol{V}, t\epsilon \boldsymbol{T^h}} P^{evt}\right) - \sum_{e\epsilon \boldsymbol{E^h}, v\in \boldsymbol{V}} c_d \cdot \left(\sum_{t\in \boldsymbol{T^h}} (A^{evt} * t) - \sum_{t\epsilon \boldsymbol{T^h}} P^{evt} \cdot ((a^e)^h + (opt^e)^h)\right) \tag{5.3}$$

Subject to:

$$P^{evt} \leq \sum_{j_{t_2}\epsilon\{j_{t_2} | (m^e{}_t, j_{t_2})\epsilon \boldsymbol{M}\}} x^v_{m^e{}_t, j_{t_2}} \qquad \forall e \in \boldsymbol{E}, \forall v \in \boldsymbol{V}, \forall t \in \boldsymbol{T} \tag{5.4}$$

$$A^{evt} \leq \sum_{i_{t_1}\epsilon\{i_{t_1} | (i_{t_1}, n^e{}_t)\epsilon \boldsymbol{M}\}} x^v_{i_{t_1}, n^e{}_t} \qquad \forall e \in \boldsymbol{E}, \forall v \in \boldsymbol{V}, \forall t \in \boldsymbol{T} \tag{5.5}$$

$$\sum_{t\epsilon \boldsymbol{T}} P^{evt} = \sum_{t\epsilon \boldsymbol{T}} A^{evt} \qquad \forall e \in \boldsymbol{E}, \forall v \in \boldsymbol{V} \tag{5.6}$$

$$\sum_{t\in \boldsymbol{T}} P^{evt} \leq 1 \qquad \forall e \in \boldsymbol{E}, \forall v \in \boldsymbol{V} \tag{5.7}$$

$$\sum_{t\in \boldsymbol{T}} A^{evt} \leq 1 \qquad \forall e \in \boldsymbol{E}, \forall v \in \boldsymbol{V} \tag{5.8}$$

$$\sum_{t\in \boldsymbol{T}} (A^{evt} \cdot t) - \sum_{t\in \boldsymbol{T}} (P^{evt} \cdot t) \geq opt^e \cdot \sum_{t\epsilon \boldsymbol{T}} P^{evt} \qquad \forall e \in \boldsymbol{E}, \forall v \in \boldsymbol{V} \tag{5.9}$$

$$\sum_{t\in \boldsymbol{T}} (A^{evt} \cdot t) - \sum_{t\in \boldsymbol{T}} (P^{evt} \cdot t) \leq lon^e \cdot \sum_{t\epsilon \boldsymbol{T}} P^{evt} \qquad \forall e \in \boldsymbol{E}, \forall v \in \boldsymbol{V} \tag{5.10}$$

$$\sum_{v\in \boldsymbol{V}, t\epsilon \boldsymbol{T}} P^{evt} \leq 1 \qquad \forall e \in \boldsymbol{E} \tag{5.11}$$

$$\sum_{e\in \boldsymbol{E}, t_1\in \boldsymbol{T}, t_1\leq t} P^{evt_1} - \sum_{e\in \boldsymbol{E}, t_2\in \boldsymbol{T}, t_2\leq t} A^{evt_2} \leq vcap \qquad \forall t \in \boldsymbol{T}, t < T, \forall v \in \boldsymbol{V} \tag{5.12}$$

$$\sum_{j_{t_2}\epsilon\{j_{t_2} | (i_{t_1}, j_{t_2})\epsilon \boldsymbol{M}\}} x^v_{i_{t_1}, j_{t_2}} \leq 1 \qquad \forall i \in \boldsymbol{I}, \forall t_1 \in \boldsymbol{T}, \forall v \in \boldsymbol{V} \tag{5.13}$$

$$\sum_{\substack{(i_{t_1}, j_{t_2})\in \boldsymbol{M} \\ t_1\leq t,\, t_2>t}} x^v_{i_{t_1}, j_{t_2}} + \sum_{i\in \boldsymbol{I}} st^v_{i_t} = 1 \qquad \forall v \in \boldsymbol{V}, \forall t \in \boldsymbol{T^h} \tag{5.14}$$

$$\sum_{l_{t_2}\epsilon\{l_{t_2} | (l_{t_2}, i_{t_1})\epsilon \boldsymbol{M}\}} x^v_{l_{t_2}, i_{t_1}} + sg^v_{i_{t-1}} = \sum_{j_{t_3}\epsilon\{j_{t_3} | (i_{t_1}, j_{t_3})\epsilon \boldsymbol{M}\}} x^v_{i_{t_1}, j_{t_3}} \qquad \forall i \in \boldsymbol{I}, \forall t_1 \in \boldsymbol{T^h}, \forall v \in \boldsymbol{V} \tag{5.15}$$

$$\sum_{\substack{j_{t_2}\epsilon\{j_{t_2}|(i_0,j_{t_2})\epsilon \boldsymbol{M}\} \\ v\in \boldsymbol{V}}} x^v_{i_0,j_{t_2}} = V \qquad i = initial\ station \tag{5.16}$$

$$F_{ij} = \left(\sum_{\substack{t_1,t_2\in \boldsymbol{T_h} \\ v\in \boldsymbol{V}}} x^v_{i_{t_1},j_{t_2}}\right)\cdot \mu \qquad \forall (i,j)\in \boldsymbol{G} \tag{5.17}$$

$$F_{ij} \le rcap_{ij} \qquad \forall (i,j)\in \boldsymbol{G} \tag{5.18}$$

$$\delta^t_{ij} = \delta^{min}_{ij} + \left(\delta^{max}_{ij} - \delta^{min}_{ij}\right)\cdot\left(\frac{F^t_{ij}}{rcap_{ij}}\right)^4 \qquad \forall (i,j)\in \boldsymbol{G}, \forall t\in \boldsymbol{T}, t<T \tag{5.19}$$

$$\delta^{t_1}_{ij} \le (t_2-t_1)\cdot x^v_{i_{t_1},j_{t_2}} + \delta^{max}_{ij}\cdot\left(1-x^v_{i_{t_1},j_{t_2}}\right) \qquad \forall (i_{t_1},j_{t_2})\in \boldsymbol{M}, \forall v\in \boldsymbol{V} \tag{5.20}$$

$$\delta^{t_1}_{ij} \ge (t_2-t_1)\cdot x^v_{i_{t_1},j_{t_2}} + \delta^{min}_{ij}\cdot\left(1-x^v_{i_{t_1},j_{t_2}}\right) \qquad \forall (i_{t_1},j_{t_2})\in \boldsymbol{M}, \forall v\in \boldsymbol{V} \tag{5.21}$$

$$t_1 + \delta^{t_1}_{ij} \le t_2 + \delta^{t_2}_{ij} \qquad \forall (i,j)\in \boldsymbol{G}, t_1,t_2\in \boldsymbol{T}, t_1,t_2<T, t_1<t_2 \tag{5.22}$$

$$\sum_{\substack{j_{t_2}\epsilon\{j_{t_2}|(i_{t_1},j_{t_2})\epsilon \boldsymbol{M}\}}} x^v_{i_{t_1},j_{t_2}} = 1 \qquad \forall v\in \boldsymbol{V}, i=(loc^v)^h, t_1=(int^v)^h \tag{5.23}$$

$$P^{evt} = 1 \qquad \forall e\in \boldsymbol{E}', v=(veh^e)^h, t=(int^v)^h \tag{5.24}$$

$$P^{evt} \in \{0,1\} \qquad \forall e\in \boldsymbol{E}, \forall v\in \boldsymbol{V}, \forall t\in \boldsymbol{T}, a^e \le t \le a^e + w^e \tag{5.25}$$

$$A^{evt} \in \{0,1\} \qquad \forall e\in \boldsymbol{E}, \forall v\in \boldsymbol{V}, \forall t\in \boldsymbol{T}, a^e + opt^e \le t \le a^e + w^e + lon^e \tag{5.26}$$

$$x^v_{i_{t_1},j_{t_2}} \in \{0,1\} \qquad \forall (i_{t_1},j_{t_2})\in \boldsymbol{M}, \forall v\in \boldsymbol{V} \tag{5.27}$$

$$F^t_{ij} \in \boldsymbol{N^0} \qquad \forall (i,j)\in \boldsymbol{G}, \forall t\in \boldsymbol{T}, t<T \tag{5.28}$$

$$\delta^t_{ij} \in \boldsymbol{N^0} \qquad \forall (i,j)\in \boldsymbol{G}, \forall t\in \boldsymbol{T}, t<T \tag{5.29}$$

Constraints (5.4)-(5.11) impose the requests' pick-up and delivery times which replace constraints (3.2)-(3.10) and (3.37)-(3.40), where variable S^{ev}_{ij} is eliminated from the model formulation. Constraints (5.12) impose that the number of passengers loaded on each vehicle during time step t to $t+1$ cannot exceed the vehicle's seating capacity. When the AT's seating capacity is larger than 1, it is possible to have ride-sharing. Constraints (5.14)-(5.16), (5.23) are updated from (4.10), (4.11), (3.15) and (4.12), eliminating the parking activities in this chapter. Constraints (5.17) compute the link flow as the total number of ATs travelling on link (i,j) within one horizon, which is an update of constraints (4.13). In this chapter, we extend the time scale of the link flow and calculate it in a cumulative way. As a result, the index t of the variables F^t_{ij} and δ^t_{ij} is eliminated, which also reduces the number of variables in $[\boldsymbol{NLIP^h}]$. Constraints (5.19)-(5.22) are updated from (3.20)-(3.23), which compute the dynamic travel time of each road link by the non-linear BPR function. Constraints (5.25)-(5.29) define the domain for the decision variables. Constraints (5.13), (5.18) and (5.24) works in the same way as (3.12), (3.17) and (4.15).

5.2.2 Lagrangian relaxation

In $[\boldsymbol{NLIP^h}]$, the vehicle routing variables $x^v_{i_{t_1},j_{t_2}}$ are related to the passenger assignment variables P^{evt} and A^{evt} by constraints (5.4) and (5.5). Such relation makes the model complicated and computationally challenging. To decouple them, we relax constraints (5.4) and (5.5) and incorporate them into the objective function (5.3) with nonnegative Lagrangian multipliers μ_1^{evt} and μ_2^{evt}. The relaxed problem can be written as follows:

$$
\begin{aligned}
[\boldsymbol{RP^h}] \quad \max Z' \\
&= \sum_{e\epsilon \boldsymbol{E^h}, v\in \boldsymbol{V}, t\epsilon \boldsymbol{T^h}} c_r \cdot P^{evt} \cdot opt^e - \sum_{\substack{(i_{t_1},j_{t_2})\in \boldsymbol{M} \\ v\in \boldsymbol{V}}} c_f \cdot x^v_{i_{t_1},j_{t_2}} \cdot d_{ij} - c_v \cdot V \\
&- \sum_{e\epsilon \boldsymbol{E^h}} c_p \cdot \left(1 - \sum_{v\in \boldsymbol{V}, t\epsilon \boldsymbol{T^h}} P^{evt}\right) \\
&- \sum_{e\epsilon \boldsymbol{E^h}, v\in \boldsymbol{V}} c_d \cdot \left(\sum_{t\in \boldsymbol{T^h}} (A^{evt} * t) - \sum_{t\epsilon \boldsymbol{T^h}} P^{evt} \cdot ((a^e)^h + (w^e)^h)\right) \\
&- \sum_{e\epsilon \boldsymbol{E^h}, v\in \boldsymbol{V}, t\epsilon \boldsymbol{T^h}} \mu_1^{evt} \cdot \left(P^{evt} - \sum_{j_{t_2}\epsilon\{j_{t_2} | (m^e{}_t, j_{t_2})\epsilon \boldsymbol{M}\}} x^v_{m^e{}_t, j_{t_2}}\right) \\
&- \sum_{e\epsilon \boldsymbol{E^h}, v\in \boldsymbol{V}, t\epsilon \boldsymbol{T^h}} \mu_2^{evt} \cdot (A^{evt} - \sum_{i_{t_1}\epsilon\{i_{t_1} | (i_{t_1}, n^e{}_t)\epsilon \boldsymbol{M}\}} x^v_{i_{t_1}, n^e{}_t})
\end{aligned}
\tag{5.30}
$$

subject to (5.6)-(5.29).

For given μ_1^{evt} and μ_2^{evt}, the optimal solution of $[\boldsymbol{RP^h}]$ provides an upper bound to the original problem $[\boldsymbol{NLIP^h}]$. The $[\boldsymbol{RP^h}]$ can be further decomposed into two sub-problems: the passenger assignment problem $[\boldsymbol{SP1}]$ and the vehicle routing problem $[\boldsymbol{SP2}]$.

The passenger assignment problem involving variables P^{evt} and A^{evt} becomes the following:

$$
\begin{aligned}
[\boldsymbol{SP1}] \quad \max Z_1 \\
&= \sum_{e\epsilon \boldsymbol{E^h}, v\in \boldsymbol{V}, t\epsilon \boldsymbol{T^h}} c_r \cdot P^{evt} \cdot opt^e - \sum_{e\epsilon \boldsymbol{E^h}} c_p \cdot \left(1 - \sum_{v\in \boldsymbol{V}, t\epsilon \boldsymbol{T^h}} P^{evt}\right) \\
&- \sum_{e\epsilon \boldsymbol{E^h}, v\in \boldsymbol{V}} c_d \cdot \left(\sum_{t\in \boldsymbol{T^h}} (A^{evt} * t) - \sum_{t\epsilon \boldsymbol{T^h}} P^{evt} \cdot ((a^e)^h + (w^e)^h)\right) \\
&- \sum_{e\epsilon \boldsymbol{E^h}, v\in \boldsymbol{V}, t\epsilon \boldsymbol{T^h}} \mu_1^{evt} \cdot P^{evt} - \sum_{e\epsilon \boldsymbol{E^h}, v\in \boldsymbol{V}, t\epsilon \boldsymbol{T^h}} \mu_2^{evt} \cdot A^{evt}
\end{aligned}
\tag{5.31}
$$

Subject to (5.6)-(5.12) and (5.24)-(5.26).

This is a linear programming formulation with binary variables P^{evt} and A^{evt}. It can be solved directly by a commercial solver.

Meanwhile, the vehicle routing problem involving variables $x^{v}_{i_{t_1},j_{t_2}}$, F_{ij} and δ_{ij} considering traffic congestion can be written as follows:

$$[\boldsymbol{SP2}] \quad \max Z_2 = -\sum_{\substack{(i_{t_1},j_{t_2})\in \boldsymbol{M} \\ v\in \boldsymbol{V}}} c_f \cdot x^{v}_{i_{t_1},j_{t_2}} \cdot d_{ij} - c_v \cdot V + \sum_{e\epsilon \boldsymbol{E^h}, v\in \boldsymbol{V}, t\epsilon \boldsymbol{T^h}} \mu_1^{evt} \cdot \sum_{j_{t_2}\epsilon\{j_{t_2}|(m^e{}_t, j_{t_2})\epsilon \boldsymbol{M}\}} x^{v}_{m^e{}_t, j_{t_2}} + \sum_{e\epsilon \boldsymbol{E^h}, v\in \boldsymbol{V}, t\epsilon \boldsymbol{T^h}} \mu_2^{evt} \cdot \sum_{i_{t_1}\epsilon\{i_{t_1}|(i_{t_1}, n^e{}_t)\epsilon \boldsymbol{M}\}} x^{v}_{i_{t_1}, n^e{}_t} \tag{5.32}$$

Subject to (5.13)-(5.23) and (5.27)-(5.29).

Sub-problem $[\boldsymbol{SP2}]$ is a non-linear optimization model due to the travel time constraints (5.19). Moreover, routing all the vehicles with dynamic travel time generates a great number of decision variables and constraints, which makes it challenging to solve. We propose an iterative assignment process to exclude the non-linear constraints (5.19) and update the link travel time based on the traffic flow that results from the optimization process, following a similar concept proposed by (Correia and van Arem, 2016a). We have mentioned in section 5.2.1 that we do not allow parking in this model, otherwise, all ATs would stay at the initial nodes and no routing would happen. This is because in [SP2] there are no demand nodes and ATs choose paths only according to the impedance of each link, which will vary with the values of the Lagrangian multipliers in several iterations. The iterative assignment process is conducted as follows:

a) Compute the initial travel times, i.e. the minimum travel time on each link as the input travel times.
b) Design the routing of each AT based on the input travel times by solving $[\boldsymbol{SP2}]$ with objective function (5.32) and constraints (5.13)-(5.18) and (5.23).
c) Calculate the traffic flow on each link according to the optimal results from step b.
d) Update the travel times according to the BPR function based on the traffic flows from step c.
e) A set of errors is computed between the updated travel times and the input travel times used in $[\boldsymbol{SP2}]$.
f) If all the errors meet the stopping criterion, then the current solution is the final solution of $[\boldsymbol{SP2}]$; otherwise, go back to step b using the updated travel times as input travel times.

In this process, the travel times do no change within the optimization model, which decreases the number of vehicle movement variables $x^{v}_{i_{t_1},j_{t_2}}$ significantly. Additionally, without constraints (5.19), $[\boldsymbol{SP2}]$ becomes a linear formulation and it is easy to solve in each iteration.

5.2.3 Upper bound and lower bound

The upper bound represents the possible best objective function value of $[\boldsymbol{NLIP^h}]$**,** which means that we will never find a feasible solution better than that one. The lower bound is the best feasible solution that has been obtained for $[\boldsymbol{NLIP^h}]$. It also indicates that the global optimal solution will not be worse than that. By solving the above two sub-problems, the summation of their objective function values with given μ_1^{evt} and μ_2^{evt} constitutes an upper bound to the $[\boldsymbol{NLIP^h}]$. However, the optimal solution of $[\boldsymbol{SP1}]$ and $[\boldsymbol{SP2}]$ may violate the relaxed constraints (5.4) and (5.5), which make the solution to $[\mathrm{OP^h}]$ infeasible. In this case, we propose a semi-optimization method to adjust the infeasible solution to a feasible one, therefore producing a lower bound to $\boldsymbol{[NLIP^h]}$**.** In each iteration of the Lagrangian relaxation algorithm, the semi-optimization method has the following steps:

a) Obtain the solution values of AT routing variable $x^{v}_{i_{t_1},j_{t_2}}$ from $[\boldsymbol{SP2}]$ as $\tilde{x}^{v}_{i_{t_1},j_{t_2}}$.
b) Use the values of $\tilde{x}^{v}_{i_{t_1},j_{t_2}}$ as an input to $[\boldsymbol{NLIP^h}]$.
c) Solve $[\boldsymbol{NLIP^h}]$ with decision variables on travel request and vehicle matching, i.e. P^{evt} and A^{evt}, objective function (5.3), and constraints (5.4)-(5.12) and (5.24)-(5.26).
d) Save the optimal solution from step c as a feasible solution to $[\boldsymbol{OP^h}]$ and obtain the lower bound of this Lagrangian iteration.

With the values of $\tilde{x}^{v}_{i_{t_1},j_{t_2}}$ from step a), the routings of all the ATs are known including the effect of traffic congestion. Nevertheless, there is no information on which requests are served by these ATs. Then we keep these routing results and use them to match the passengers' requests in the dimensions of space and time by solving the model in step c. If there are some requests satisfied by the ATs, then this is a feasible solution of providing the AT service.

After each Lagrangian iteration, the multipliers μ_1^{evt} and μ_2^{evt} will be updated using the standard sub-gradient procedure (Fisher, 1981) as follows:

$$(\mu_1^{evt})^{l+1} = \max\{0, (\mu_1^{evt})^{l} + \rho^{l} \cdot \left(\tilde{P}^{evt} - \sum_{j_{t_2} \epsilon \{ j_{t_2} | (m^{e}{}_{t}, j_{t_2}) \epsilon \boldsymbol{M} \}} \tilde{x}^{v}_{m^{e}{}_{t}, j_{t_2}} \right) \} \quad \forall e \in \boldsymbol{E}, \forall v \in \boldsymbol{V}, \forall t \in \boldsymbol{T^h} \tag{5.33}$$

$$(\mu_2^{evt})^{l+1} = \max\{0, (\mu_2^{evt})^{l} + \rho^{l} \cdot \left(\tilde{A}^{evt} - \sum_{i_{t_1} \epsilon \{ i_{t_1} | (i_{t_1}, n^{e}{}_{t}) \epsilon \boldsymbol{M} \}} \tilde{x}^{v}_{i_{t_1}, n^{e}{}_{t}} \right) \} \quad \forall e \in \boldsymbol{E}, \forall v \in \boldsymbol{V}, \forall t \in \boldsymbol{T^h} \tag{5.34}$$

$$\rho^{l} = \frac{\pi^{l} \cdot (UB^{l} - LB^{*})}{\sum_{e \epsilon \boldsymbol{E}, v \in \boldsymbol{V}, t \epsilon \boldsymbol{T^h}} \left(\tilde{P}^{evt} - \sum_{j_{t_2} \epsilon \{ j_{t_2} | (m^{e}{}_{t}, j_{t_2}) \epsilon \boldsymbol{M} \}} \tilde{x}^{v}_{m^{e}{}_{t}, j_{t_2}} \right)^2 + \sum_{e \epsilon \boldsymbol{E}, v \in \boldsymbol{V}, t \epsilon \boldsymbol{T^h}} \left(\tilde{A}^{evt} - \sum_{i_{t_1} \epsilon \{ i_{t_1} | (i_{t_1}, n^{e}{}_{t}) \epsilon \boldsymbol{M} \}} \tilde{x}^{v}_{i_{t_1}, n^{e}{}_{t}} \right)^2} \tag{5.35}$$

where $\tilde{P}^{evt}$, $\tilde{A}^{evt}$ and $\tilde{x}^{v}_{i_{t_1},j_{t_2}}$ are the values of decision variables from [$\boldsymbol{SP1}$] and [$\boldsymbol{SP2}$] in the lth iteration; $\pi^{l=1}$=2 as an initial value and it will be halved when the upper bound has failed to improve in 3 Lagrangian iterations. The initial values of the Lagrangian multipliers are set as μ_1^{evt}=1.5, μ_2^{evt}=1.5, $\forall e \in \boldsymbol{E^h}, \forall v \in \boldsymbol{V}, \forall t \in \boldsymbol{T^h}, \forall h \in \boldsymbol{H}$.

The following pseudo-code shows the customized Lagrangian relaxation algorithm, where l is the Lagrangian iteration number and k is the traffic assignment iteration number. This solving process is used to solve [$\boldsymbol{NLIP^h}$] for horizon h, which should be embedded in the rolling horizon framework. It replaces Step 2 in the pseudo-code presented in Section 4.

Step 0: Initialize $(\mu_1^{evt})^{l=1} = 1.5, (\mu_2^{evt})^{l=1} = 1.5, \forall e \in \boldsymbol{E^h}, \forall v \in \boldsymbol{V}, \forall t \in \boldsymbol{T^h}$

Step 1: Solve [$\boldsymbol{SP1}$] with $(\mu_1^{evt})^l$ and $(\mu_2^{evt})^l$;

Obtain the value of $(Z_1)^l$, $\tilde{P}^{evt}$ and $\tilde{A}^{evt}$

Step 2: Solve [$\boldsymbol{SP2}$]

Step 2.0: Initialize $(\tilde{\delta}_{ij})^{k=0}$=$\delta_{ij}^{min}$

Step 2.1: Solve [$\boldsymbol{SP2}$] with $(\tilde{\delta}_{ij})^k$ as input values instead of decision variables;

save the value of $\tilde{F}_{ij}$ from variable F_{ij}

Step 2.2: If $k = 0$ then

$$\left(Vol_{ij}\right)^0 = \tilde{F}_{ij} \quad \forall (i,j) \in \boldsymbol{G}$$

else

$$\left(Vol_{ij}\right)^k = \left(1 - \frac{1}{K}\right) \cdot \left(Vol_{ij}\right)^{k-1} + \frac{1}{K} \cdot \tilde{F}_{ij} \quad \forall (i,j) \in \boldsymbol{G}$$

end-if

Step 2.3: Update the link travel times

$$(\tilde{\delta}_{ij})^{k+1} = \delta_{ij}^{min} + \left(\delta_{ij}^{max} - \delta_{ij}^{min}\right) \cdot \left(\frac{\left(Vol_{ij}\right)^k}{rcap_{ij}}\right)^4 \quad \forall (i,j) \in \boldsymbol{G}$$

Step 2.4: If $(\tilde{\delta}_{ij})^{k+1} - \left(\tilde{\delta}_{ij}\right)^k \leq stopping\ criterion\ 1 \quad \forall (i,j) \in \boldsymbol{G}$ then

save the values of $(Z_2)^l$ and $\tilde{x}^{v}_{i_{t_1},j_{t_2}}$

go to Step 3

otherwise, $k = k + 1$, go to Step 2.1

Step 3: Find the feasible solution

Step 3.1: Put the values of $\tilde{x}^{v}_{i_{t_1},j_{t_2}}$ from Step 2.4 into [$\boldsymbol{NLIP^h}$]

Step 3.2: Solve [$\boldsymbol{NLIP^h}$]

Step 3.3: Save the values of variable P^{evt} and A^{evt} together with $\tilde{x}^{v}_{i_{t_1},j_{t_2}}$ from Step 2.4,

then this is a set of feasible solution to [$\boldsymbol{NLIP^h}$]

Step 3.4: Obtain the value of $\left(Z^h\right)^l$

Step 4: $UB^l = (Z_1)^l + (Z_2)^l$

If $h = 1$ then $LB^* = \left(Z^h\right)^l$

else if $LB^* < \left(Z^h\right)^l$ then $LB^* = \left(Z^h\right)^l$

end-if

Step 5: $Gap = (UB^l - LB^*)/\ UB^l \cdot 100\%$
$IfGap \leq stopping\ criterion$ 2 then finish;
otherwise, go to Step 6
Step 6: Update Lagrangian multipliers according to (5.33)-(5.35);
$h = h + 1$, go to Step 1

5.3 Case study

The model is applied to the same case study in chapter 4. In this section we present some changes and new information for the case study application.

5.3.1 Application set-up

According to the updated assumptions and settings of the AT system, some of the parameters are modified as follows.

- c_v=20 euro/vehicle/day.
- c_p=1 euro/request.
- c_d=0.2 euro/min.

The seating capacity of the ATs is set as 2, which indicates that an AT can at most take two shared ride passengers at the same time. The maximum waiting time for each request is 22.5 minutes. We consider a daytime operation of 15 hours from 7:00-22:00 for the AT service, meaning that the rolling-horizon framework contains 60 horizons. For this application, 22,240 requests from section 4.3.2 are all considered as real-time requests in this chapter. We set K=10 meaning that for each running process of [$\boldsymbol{SP2}$] 10+1 iterations are allowed, working with the stopping criterion that the error for each link travel time should be no more than 10%.

5.3.2 Base scenario

We programmed the model in Mosel language and solved the problem with the FICO® Xpress Solver in an i5 processor @3.10GHz, 12.00 GB RMA computer under a Windows 7 64-bit operating system. FICO® Xpress Solver provides optimization algorithms and technologies to solve linear, mixed integer and non-linear problems. The scenario with 500 ATs and a seating capacity of 2 is set as the base scenario.

The model [$\boldsymbol{NLIP^h}$] is firstly run for one horizon (h=48) as an example to illustrate how it works the type of results that can be obtained. The 48th horizon is chosen because there are 400 travel requests which is close to the average number of requests per horizon. We run the model from the first horizon to the 47th horizon and use the vehicle status and demand satisfaction information from the 47th horizon to guarantee the continuity. Then model [$\boldsymbol{NLIP^h}$] is solved for the 48th horizon with 300 Lagrangian iterations.

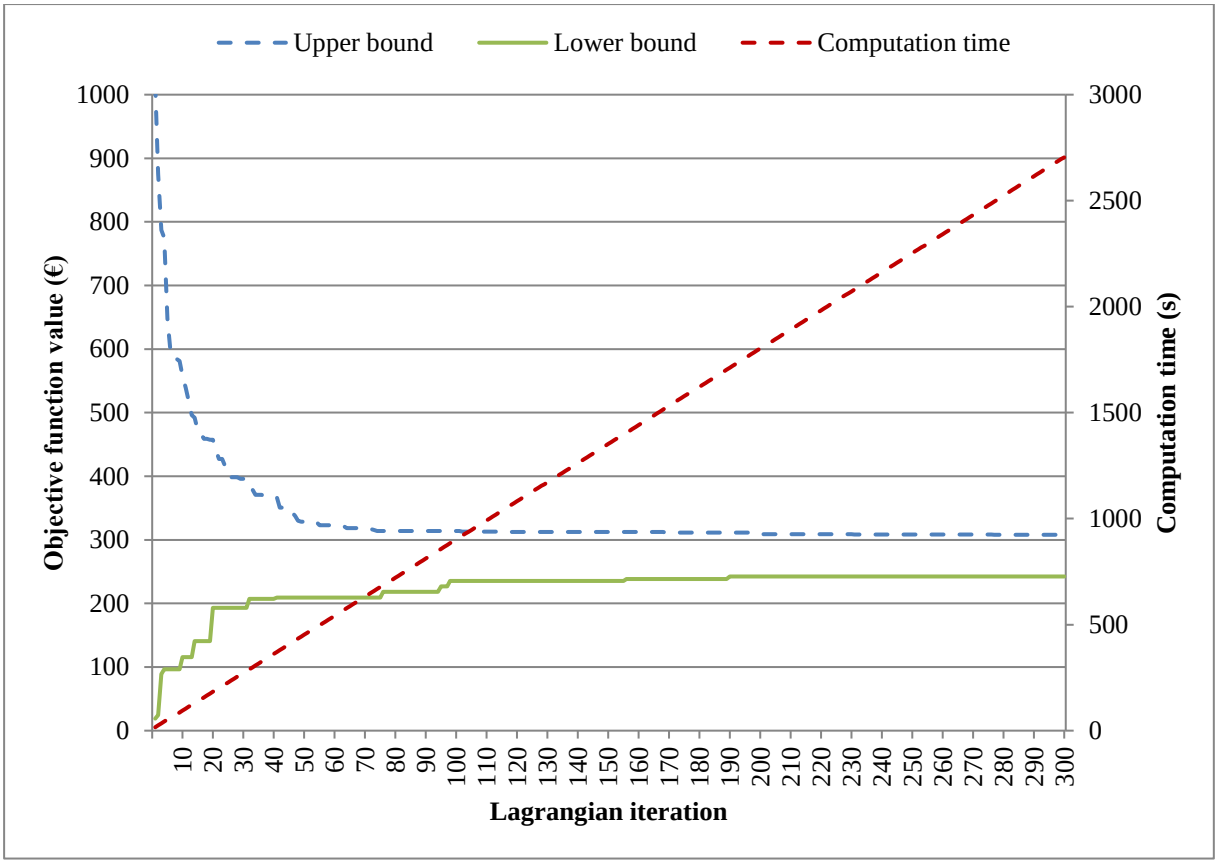


Figure 5.3 Computation results for one horizon

The graphical and numerical results are shown in Figure 5.3 and Table 5.2. The upper bound of [$\boldsymbol{NLIP^h}$] drops rapidly within the first 50 iterations from about 1000 to 328.4, while the lower bound also has an obvious growth at the same time from about 10 to 209.4. The gap between the current best solution and the current best bound is about 36%. From the 51st iteration to the 100th, the model begins to converge at a stable speed and the upper bound and the lower bound get closer, which narrows the gap to 25%. After iteration 100, the convergence slows down but the upper bound and the lower bound keep improving. The lower bounds from the 100th iteration to the 150th iteration are all the same meaning that during this time the model cannot find a better feasible solution to the [$\boldsymbol{NLIP^h}$] to replace the current one. However, the upper bound keeps going down and this contributes to closing the gap. After 190th iteration, the lower bound does not change until this solving process reaches 300 iterations, but the upper bound slightly increases, with the final gap being 21%.

The computation time is proportional to the number of Lagrangian iterations. With more iterations, it has a higher probability to find a better feasible solution and a lower upper bound, which is able to show that this solution is closer to the optimal one. However, more iterations also bring longer computation time. In this one horizon example, iteration 50,100 and 200 are reasonable breakpoints: the gaps are 36%, 25% and 22% respectively, showing that the quality of the solutions is acceptable. Therefore, we decide to choose L=50,100 and 200 as the maximum number of Lagrangian iterations to run all the horizons and compare the final

results of the base scenario. The computation results with different Lagrangian iterations are presented in Table 5.3.

Table 5.2 Computation results for one horizon

Until iteration	**10**	**20**	**30**	**40**	**50**	**60**	**70**	**80**	**90**	**100**
Upper bound	558.1	457.3	396.2	368.3	328.4	321.4	316.8	313.8	313.8	313.8
Lower bound	115.4	192.9	192.9	207.4	209.4	209.4	209.4	218.4	218.4	235.4
Gap	79%	58%	51%	44%	36%	35%	34%	30%	30%	25%
Until iteration	**110**	**120**	**130**	**140**	**150**	**160**	**170**	**180**	**190**	**200**
Upper bound	312.9	312.3	312.3	312.3	312.3	312.2	311.7	311.7	311.5	308.9
Lower bound	235.4	235.4	235.4	235.4	235.4	238.4	238.4	238.4	242.4	242.4
Gap	25%	25%	25%	25%	25%	24%	24%	24%	22%	22%
Until iteration	**210**	**220**	**230**	**240**	**250**	**260**	**270**	**280**	**290**	**300**
Upper bound	308.9	308.8	308.8	308.6	308.5	308.2	308.2	308.0	308.0	308.0
Lower bound	242.4	242.4	242.4	242.4	242.4	242.4	242.4	242.4	242.4	242.4
Gap	22%	21%	21%	21%	21%	21%	21%	21%	21%	21%

When 50 Lagrangian iterations are used for each horizon, the model spends 15.9 min per horizon to obtain a feasible solution with the average gap of 32.4% and the median value of the horizon gap 29.5%. While if we increase the iteration number to 100, the model spends 34.3 min per horizon to find an acceptable feasible solution. At the same time, the average gap per horizon is decreased from 32.4% to 26.8%, and the median of the gaps is also decreased from 29.5% to 23.5%. This demonstrates that increasing the number of Lagrangian iterations helps the model to close the gap, while along with the computation time growing. When 200 Lagrangian iterations are applied, the average gap keeps decrease to 23.0%, while the median of the horizon gaps has a more significant drop from 23.5% to 9.0%. This means that half of the horizons have the optimization gap below or equal to 9.0% and among the rest half horizons, some of them are difficult to solve which makes the average gap higher than the median.

Table 5.3 Results for the base scenario with 50, 100 and 200 Lagrangian iterations

Number of iterations	Average computation time per horizon (min)	Average gap per horizon	Median of the horizon gaps
50	15.9	32.4%	29.5%
100	34.3	26.8%	23.5%
200	57.6	23.0%	9.0%

In the following experiments, we believe that iteration 200 is an acceptable stopping criterion to obtain a feasible solution with good quality. Additionally, the iterative process will stop when the gap is lower or equal to 1% even if the number of iterations is lower than 200, which is another stopping criterion. Even though the computation time of this model is not fast enough to handle a real-time problem, we still can expect that the advances of computer technologies and new algorithms can accelerate significantly the solving of this model in the future.

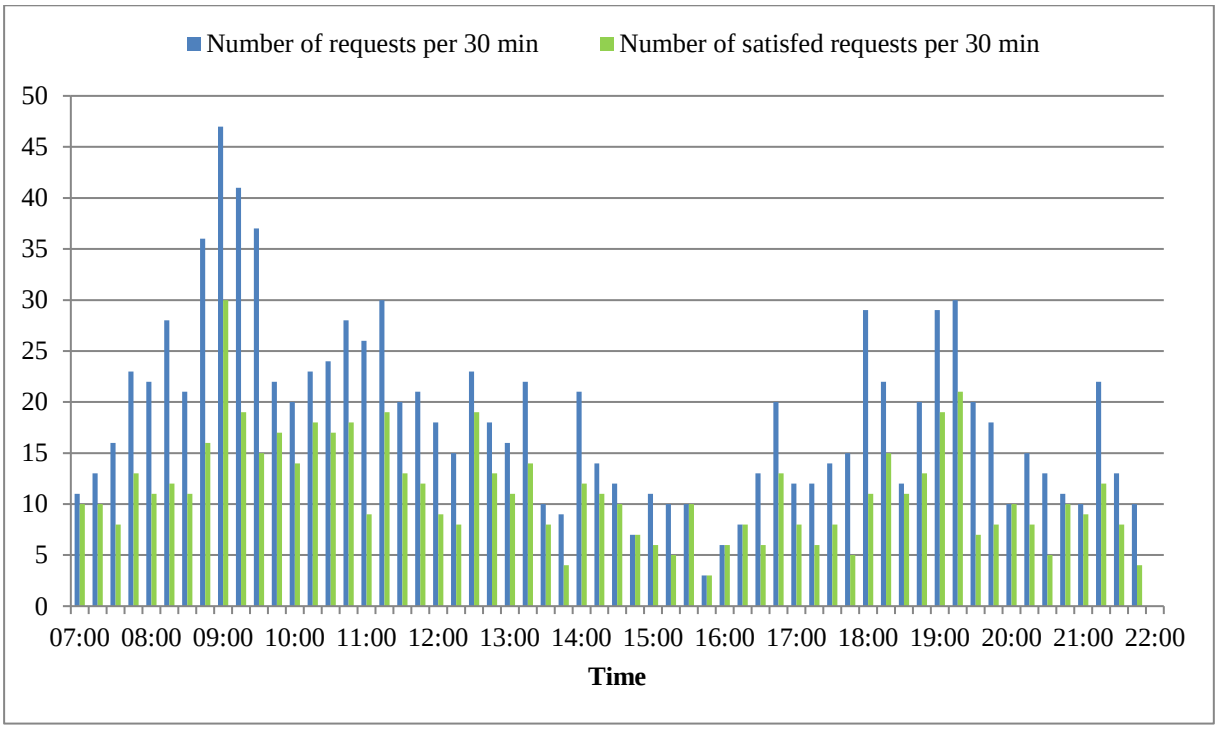


Figure 5.4 Requests distribution

The time distribution of the satisfied requests using 200 Lagrangian iterations is shown in Figure 5.4. The demand is not uniformly distributed during the day, which creates the "peak hours" and the "off-peak hours" in the rolling horizon framework. In this scenario, 13460 passengers in total are served out of 22240, meaning that 60.5% of the requests are served by ATs. In the morning peak hours, the AT system has a relatively low satisfaction rate. The afternoon peak hours have a similar tendency. However, the system provides a higher share of ATs for the low-demand time periods and some of them have a 100% satisfaction rate. This figure illustrates that the high demand for the morning and afternoon peak hours exceeds the transport capacity of the system. It also means that in the off-peak horizons, 500 ATs are redundant and generate idle time. Moreover, the objective function of [$\boldsymbol{NLIP^h}$] is to maximize the daily profit of the AT company, which means that the system will select the requests to be served when they bring more profit. The objective function also considers the service quality by adding the penalty for rejected requests and the delivery delay. Therefore, the AT system aims at having a higher AT share to gain more revenue and keep a higher service quality to avoid more penalties. Section 6.3 and 6.4 presents more detailed discussions on this.

5.4 Experiments and results

To test the performance of the proposed model, we vary the scenarios in two dimensions: the number of ATs and the seating capacity of each AT. We also vary the values of AT price rate, rejection penalty and delay penalty to conduct a sensitivity analysis with respect to the cost components. The descriptions for all the scenarios are shown in Table 5.4 and the optimization results can be seen in Table 5.5. AT share is defined as the number of requests ATs are serving divided by the total number of requests. "Percentage of time in service" is calculated by dividing the total travel time when the ATs are transporting passengers by the total time for the whole operation period. The delay is computed as the time spent by each traveller from the desired departure time to the real arrival time minus the shortest travel time. It includes the departure delay (waiting time) and the congestion delay (real travel time minus the shortest travel time).

Table 5.4 Scenario description

Scenario	Description	V	$vcap$	c_r euro /min	c_p euro /request	c_d euro /min
0	Base	500	2	1	1	0.2
1	Fewer ATs in the system	400	2	1	1	0.2
2	More ATs in the system	600	2	1	1	0.2
3	Lower seating capacity (no ride-sharing)	500	1	1	1	0.2
4	Higher seating capacity for ride-sharing	500	3	1	1	0.2
5	Lower price rate	500	2	0.5	1	0.2
6	Higher price rate	500	2	1.5	1	0.2
7	No penalty for request rejection	500	2	1	0	0.2
8	Higher penalty for request rejection	500	2	1	2	0.2
9	No penalty for delivery delay	500	2	1	1	0.0
10	Higher penalty for delivery delay	500	2	1	1	0.4

5.4.1 Fleet size variation

Scenarios 0,1 and 2 are tested with different fleet sizes: 400, 500 and 600 (Table 5.5). When there are only 400 ATs in the city in scenario 1, the system produces 84.2×103 euros as daily profit, which is 6% lower than the base scenario with 500 ATs, while with 600 ATs the system has 1% higher daily profit. This is because of the different components considered in the system's profit. When more ATs are available for the passengers, they will serve more passengers' requests and the AT company will get more revenue from them, which can be

seen from the columns of “Total revenue” and “Number of satisfied requests”. At the same time, the penalty paid for rejecting requests will be lower, which also contributes to the system profitability. However, the AT’s depreciation cost and the fuel cost also increase with having more vehicles in the system, since these costs are almost proportional to the fleet size.

In these scenarios, the fleet size is a critical parameter with respect to the AT share in urban mobility. When there are only 400 vehicles in the taxi system, 12220 requests are satisfied in total. In this case, the AT share, which can also be considered as the satisfied rate, is the lowest among all the scenarios. When the fleet size increases to 500, the AT share increases by 5.6%. This trend continues in the next scenario with 600 ATs, but the growth rate is falling. This means that with the second 100 vehicles added the system cannot provide a significant growth on service coverage, which also means that from that moment on, other indicators become critical in respect to the improvement of daily profit, i.e. the service quality.

With respect to vehicle usage, the most efficient scenario is the one with 400 ATs, with 30.6 requests per vehicle. This can also be seen in the column “Average travel time in service”. The percentage of time in service is 52.4%, which is the highest of the three. The 600 ATs scenario has the lowest percentage of time in service (41.3%), which means that 58.7% of the total operation time ATs are idle. The added ATs are redundant in off-peak horizons, while in the peak horizons they are more helpful and can contribute to a higher AT share. On average, each passenger spends 14.3 euros per trip and stays 15.4 min inside the vehicle in the 500 vehicle scenario. These values almost the same for scenario 1 and 2, meaning that the fleet size does not have a significant influence on these two indicators.

The average delay per satisfied request keeps decreasing from the fewer-ATs scenario to the more-ATs one. This happens because of the different average waiting time for each scenario and the average congestion time. The variation of the fleet size has a significant influence on the waiting time of the satisfied requests for the scenarios of 400 and 500 ATs, with 0.7 min reduction. While from 500 ATs to 600 ATs, it becomes only 0.1 min. This demonstrates that the first additional 100 ATs helps a lot regarding the waiting time and accomplish more requests to increase service performance. In this case, the AT’s service quality is improved from the passengers’ perspective, which is consistent with the setting of the objective function (5.3). However, the second additional 100 ATs does not contribute a lot to the reduction in the waiting time for departure, but they indeed satisfy more requests which increase the AT share to 64.6%.

5.4.2 Ride-sharing variation

In these experiments, we apply the flat AT price rate for the different seating capacities, in order to analyse the interaction between the level of ride-sharing and the system profitability, the AT share and the service quality.

Table 5.5 Optimization results for the reference scenarios

Scenario	Profit per day ($\times 10^3$ euros)	Total revenue ($\times 10^3$ euros)	Total fuel cost ($\times 10^3$ euros)	Total penalty cost for rejected requests ($\times 10^3$ euros)	Total depreciation cost ($\times 10^3$ euros)	Total penalty cost for delay ($\times 10^3$ euros)	Number of satisfied requests	AT share	Avg.* satisfied requests Per AT	Avg. travel time in service per AT (min)	Percentage of time in service	Avg. price per satisfied request (euros)	Avg.travel time per satisfied request (min)	Avg. delay per satisfied request (min)	Avg. waiting time per satisfied request (min)	Avg. congestion delay per satisfied request (min)
0) Base	**89.6**	**192.1**	**45.1**	**8.8**	**10.0**	**38.6**	**13460**	**60.5%**	**26.9**	**415.1**	**46.1%**	**14.3**	**15.4**	**14.3**	**13.0**	**1.3**
1) Fewer ATs	84.2	174.8	36.0	10.0	8.0	36.6	12220	54.9%	30.6	471.5	52.4%	14.3	15.4	15.0	13.7	1.3
2) More ATs	90.4	205.1	54.0	7.9	12.0	40.8	14360	64.6%	23.9	371.3	41.3%	14.3	15.5	14.2	12.8	1.4
3) Lower seating capacity	43.7	140.3	45.1	12.5	10.0	29.0	9700	43.6%	19.4	299.0	33.2%	14.5	15.4	14.9	14.0	0.9
4) Higher seating capacity	118.4	224.0	44.8	6.7	10.0	44.1	15560	70.0%	31.1	489.1	54.3%	14.4	15.7	14.2	12.9	1.3
5) Lower price rate	-7.0	91.0	45.3	9.7	10.0	33.0	12560	56.5%	25.1	388.6	43.2%	7.2	15.5	13.1	12.1	1.0
6) Higher price rate	151.1	255.4	44.8	8.3	10.0	41.2	13980	62.9%	28.0	430.8	47.9%	18.3	15.4	14.7	13.4	1.3
7) No pen.** for rejection	91.4	181.2	45.1	0.0	10.0	34.7	12480	56.1%	25.0	394.8	43.9%	14.5	15.8	13.9	12.6	1.3
8) Higher pen. for rejection	83.7	195.5	45.0	17.2	10.0	39.6	13640	61.3%	27.3	419.0	46.6%	14.3	15.4	14.5	13.2	1.3
9) No pen. for delay	139.2	202.5	45.3	8.0	10.0	0.0	14280	64.2%	28.6	519.3	57.7%	14.2	18.2	20.4	16.4	4.0
10) Higher pen. for delay	28.5	146.3	44.7	11.9	10.0	51.2	10300	46.3%	20.6	310.5	34.5%	14.2	15.1	12.4	11.6	0.8

* Average
** Penalty

The seating capacity indicates the maximum number of passengers allowed to share a ride at the same time. When it varies from 1 to 3, it also increases the total transport capacity of the AT system. It is obvious from Table 5.5 that the improvement of the system profitability is significant when the system changes its serving scheme from individual AT service to ride-sharing AT service. When the seating capacity increases from 2 to 3, the improvement of the total profit happens but not with the same magnitude as when it goes from 1 to 2. The AT service company earns 45.9×103 euros more in allowing two passengers to share a ride, compared with no ride-sharing. Looking at the constituent parts of the profit, the increase in the total revenue is the main contributor to the profit. This is because when the seating capacity of the ATs is doubled, the system is able to serve more travel requests, which can be seen from the column "Number of satisfied requests" and "AT share". On one hand, it will bring more revenue to the AT company; while on the other hand, it reduces the penalty paid to the rejected passengers who have to use the public transport as an alternative to accomplish their trips. However, when three passengers are allowed to share a ride the AT system serves 2100 more requests which provide 31.9×103 euros more to the system revenue, 2.1×103 euros less on the rejection penalty cost and 5.5×103 euros less on the delay penalty. This improvement is not as considerable as the one obtained going from no ride-sharing to two seat sharing. This means that the system is not able to serve more requests, since other reasons become a constraint for the demand satisfaction, e.g. the departure and arrival time windows.

At the same time, the usage of the vehicles is also improved. The two-seat scenario has 7.5 more requests served per AT and 116.1 min more in service than the one-seat scenario. With the doubled value of seating capacity, these values are not as doubled as is could be expected. This happens because there are still some other constraints for satisfying the requests, e.g. hard constraints like the request's individual time window and soft constraints like the delay penalty due to the ride-sharing. This also means that the ATs are not always taking shared ride passengers. The trend also happens in the scenario with 3 ride-sharing passengers, but the increases are not as significant as the previous one.

Due to the sufficient transporting capacity, the average delay drops from 14.9 min to 14.3 min and then to 14.2 min, which indicates the improvement of the service level offered to clients. To be more specific, the waiting time keeps going down among the three scenarios; while the congestion delay has an exceptional case: 1.3 min for scenario 0 and 4, which is 0.4 min higher than scenario 3. Even though, the AT service quality regarding the delay is improved because of the ride-sharing.

5.4.3 Sensitivity analysis on price rate

The price rate has a critical impact on the AT system's daily profit. It is easy to see that the number of satisfied requests by the ATs increases from 12560 to 13460 and then to 13980 when increasing the price rate from 0.5 euro/min to 1 euro/min then to 1.5 euro/min. The same trend also happens with the "Total penalty cost for rejected requests" and "Total penalty

cost for delay", showing that the monetary benefits gaining from the increase of the AT share remain at a relatively low level. The reason for this is that when the system can have more revenue from the same requests by increasing the price rate, the difference between the revenue and the cost for each satisfied request is more obvious. So the system has a greater motivation to satisfy more requests. At the same time, more requests may lead to a more congested road network, which increases the penalty paid for the travel delay. So the system should find a balance between the profit benefit from the higher AT share and the profit loss from the delivery delay.

Since the fleet size is a constant for these three scenarios, the fuel cost and the vehicle depreciation costs have almost the same values in the objective function. Therefore, the difference between daily profits results primarily from the revenue earned from AT passengers. When the price is 0.5 euros/min, the AT company can obtain 91.0×103 euros as revenue; while this value is almost doubled to 192.1×103 euros when the price increases to 1 euro/min. Similarly, the revenue of the 1.5 euros/min scenario is almost three times the one obtained with 0.5 euros/min. This means that the system revenue is approximately proportional to the price rate of the AT service. This can also be seen in the column "Avg. price per satisfied request", where the average monetary cost for one request is doubled from scenario 5 to scenario 0. While in scenario 6 the average price per request is lower than three times of that price in scenario 5, indicating that when the price is higher, the system tends to serve more shorter (cheaper) requests. However, the revenue for scenario 5 is not high enough to cover the other costs for the system, leading to a negative profit as the final result.

5.4.4 Sensitivity analysis on rejection penalty

The rejection penalty is charged when a potential request is rejected by the AT system and the value of it is set according to the cost of a public transport alternative. This penalty is aimed at improving service performance and encourage the system to take more passengers. It also allows the AT system to make the decision that when some requests are either unprofitable or un-satisfiable, it can reject these requests.

The rejection penalty brings profit loss to the system. The profit decreases 1.8 ×103 euros when paying 1 euro for each rejected request when compared to not having a penalty. This is due to the rejection penalty of 8.8×103 euros, the additional 3.9×103 euros for the delay penalty and the extra revenue of 10.9×103 euros. If the rejection penalty increases to 2 euros per unsatisfied request, it brings 3.4×103 euros more to the revenue, 8.4×103 euros more to the rejection penalty cost and 1.0×103 euros more to the delay cost, which in total contributes 5.9×103 euros less to the daily profit of the AT system.

The rejection penalty plays a role in promoting the growth of the AT share, increasing from 12480 to 13460 requests, and from 13460 to 13640 requests. The revenues of these served requests have slight growth for these three scenarios. However, the average price per satisfied request of the scenario with no penalty for rejection is the highest, indicating that the average

shortest travel times for the satisfied requests in scenarios 0 and 8 are shorter than scenario 7. This happens because the system tends to satisfy requests that have shorter travel times to avoid lower rejection costs. But the system profit still cannot benefit from this penalty policy, since it is not helpful in increasing the revenue from the passengers.

The travel delay firstly increases from 13.9 min to 14.3 min, then decreases to 14.5 min, for scenario 7, 0 and 8. This trend also happens to the "Average waiting time per satisfied request" and the "Average congestion delay per satisfied request". The first increase from no penalty to one euro is due to the growth in the total delay penalty and the average satisfied requests per AT because the ATs are busier to satisfy more requests and this leads to more delay for each request. A small increase happens as a result of increasing one euro to two euros the rejection penalty, indicating that it is better to serve more requests but with a good service quality to avoid the costs for the added delay. In this way, the daily profit of the system is maintained at a specific level.

5.4.5 Sensitivity analysis on the delay penalty

The daily profit is sensitive to the delay penalty. When there is no charge for late arrival, the system only needs to guarantee each satisfied request to be picked up and dropped off within its allowed departure and arrival time window. However, if the AT company needs to pay for the time delay of delivering the passenger, the system not only needs to comply with the time windows but it also needs to provide arrivals that happen as early as possible. This penalty reflects the service quality of the AT system.

For the scenario with no delay penalty, the system has the highest daily profit (139.2×103 euros), the highest AT share (64.2%) and the highest average delay (4.0 min) in transporting passengers among scenario 9,0 and 10. This is the most profitable scenario among the three due to the high revenues from the satisfied requests and the zero penalty cost for the time delay. However, the other two scenarios have 820 and 3980 fewer requests being satisfied by ATs, which generate lower revenue for the system. This also makes the system pay the lowest value for the rejection penalty when there is no delay penalty. Regarding the vehicle usage, the system becomes less efficient when increasing the value of delay penalty, which can be seen in the "Average satisfied requests per AT" and the "Percentage of time in service". This is not surprising since the same number of ATs are used to provide a different number of satisfied requests. The time delay has a sharp decrease (from 20.4 min to 12.4 min) with the increase of the delay penalty. This tendency also occurs in the waiting time and the congestion delay. With the high penalty for the time delay, the system profitability is quite sensitive to the late arrival. In some extreme cases, if a request is delivered with a long delay time, it may happen that the delay penalty is higher than the revenue obtained from this request, which means satisfying this passenger will be damaging for the profit. Under these circumstances, the system decides to serve fewer requests guaranteeing the service level (short delay time) to protect the profit. Therefore, the delay penalty is an appropriate control parameter to assure the service quality being provided by the ATs.

5.5 Conclusions

AVs have drawn great attention in the recent decades evolving from just a concept to reality with several pilot studies. An emphasis has been put on technology in recent research focusing on transport reliability and safety, and the interactions between the AVs and the other vehicles, pedestrians and cyclists. Nevertheless, some questions related to the application of AVs have hardly been addressed, such as their use as public transport.

This chapter proposed a NLIP model [$\boldsymbol{NLIP^h}$] to study the DARP of ATs under dynamic travel times created by the ATs themselves. The model has as its main objective to maximize the total daily profit of such system by deciding on each AT's routing with real-time information. The penalties of the rejected requests and the delivery delay are considered in the objective function to guarantee the performance of the AT service. We argue that this type of model is needed in future situations in which a large number of requests and a large number of vehicles are involved. Therefore, the travel time of the ATs has been modelled as a function of the traffic flow of the ATs themselves on the road links in the constraints. The model considers the demand as real-time requests which are managed via a rolling horizon structure thus allowing the vehicles' routing being optimized horizon by horizon. Even though this framework makes it possible to decrease the scale of the problem in each optimization process, the proposed model still involves a large number of integer variables, which makes it an NP-hard one. We develop a customized Lagrangian relaxation algorithm to solve the mathematical model, which decomposes the [$\mathrm{OP^h}$] into two sub-problems. Based on the solutions gained from the two sub-problems, we are able to find the best bound of the original problem and the current best solution, which is feasible to the original problem. In this way, it is not only possible to find a near optimal solution but also have a notion of how good this solution is.

The model and the solving algorithm was applied to the case study city of Delft, The Netherlands, with 15 hours service period and 22,240 travel requests generating from the road network with 46 nodes and 66 links. From that application, it was possible to make several conclusions.

The customized Lagrangian solving algorithm is able to solve the proposed NP-hard problem and obtain a good quality feasible solution within the acceptable computation time. The solving times are still high enough to forbid their application in real time systems but we believe that this is an initial version of the type of algorithms that must be built and run in the future. The fleet size is an important factor in system profitability and the satisfaction of the requests. When the fleet increases, the number of satisfied requests are not increasing with the same rate, which demonstrates that the fleet size is not the only reason to constrain the satisfaction of the requests. In fact, the uneven time distribution of the demand and the strict departure and arrival time windows are all important factors in influencing the AT share. The model shows that the AT passengers will experience a delivery delay when considering the traffic congestion generated by the AT themselves. With ride-sharing, the AT system has

more capacity to provide better service regarding the number of satisfied requests and the average delay time. The system profitability is sensitive to the price rate of the AT service. The different values of rejection penalty will lead to changes in the ATs' daily profit and the number of satisfied requests. The delay penalty seems to be a proper control parameter to guarantee a certain service quality being offered by the ATs systems. When this penalty is not considered, the AT system will have a higher AT share, along with a longer delivery delay for the satisfied requests.

The current version of the optimization model is not completely practice-ready, due to the computation time and the computation gap between the upper and lower bounds. Nevertheless, this formulation can be the basis for future work. It can be used to develop heuristics for the ATs' DARP problem, which may accelerate the optimization process. In addition, more efforts can be devoted to improving the solution algorithm and close the computation gap of the proposed model. In this research, the AT system covers 50%-70% of the total transport demand, with the objective to maximize the AT company's profit. Future research could select the share of ATs as an objective and investigate the changes in performance. There is also the possibility of involving choice modelling to analyse people's preferences for using ATs. This would allow running a sensitivity analysis on the influence of the AT price rate on the potential demand and the overall system profitability.

Chapter 6

Conclusions and recommendations

This thesis has been mainly motivated by the challenges in the planning and the operation of future AT systems outlined in Chapter 1. The main objective was to develop an optimization framework for using ATs in a public transport system and test the performance and profitability of the proposed system. Several research questions, which were stated under the main research objective, were answered throughout Chapters 2 to 5. This final chapter discusses the main research findings, conclusions, implications for practice and points for possible future research directions. Section 6.1 summarizes the main findings, limitations and conclusions regarding the planning and operation of ATs. Section 6.2 highlights the contributions of this thesis. Section 6.3 discusses the implications that can be extracted from this thesis in what respects to the possible application of the AT systems in practice. Section 6.4 sets paths for future research directions.

6.1 Main findings, limitations and conclusions

The objective of this thesis was to establish an optimization framework to address the planning and operational strategies that AT systems should follow in order to serve urban mobility demand. It aimed at answering the main research question: *How should an AT system be designed in order to optimally serve people's urban travel demand?*

To answer this question, research was divided into four parts. Here we structure the main findings on answering the four key sub-research questions defined in Section 1.2.

Research question 1: How should the service area of ATs be designed for the last mile transport of train trips?

To answer this research question, we considered two service schemes for a future AT system: free service (S1) and full service (S2). In the first, the company may select operational zones and also the trips that it prefers to satisfy, whilst in the second, when selecting a zone to be part of the operational area, all trips with origin and destination at that zone must be satisfied. The daily profit has been selected as the indicator for the AT's performance and the optimization objective is naturally to maximize this indicator. The models were applied to the case study of Delft Zuid train station in Delft, the Netherlands. This case study provided insights into the effect of service zone location and trip selection on the profitability of a future AT system.

Using ATs can significantly improve the daily profit of such taxi company comparing to human-driven cars. In model S1, with a small fleet, the staff cost would not influence the system profit significantly comparing to an AT system. However, with bigger fleets, human-driven cars would lead to fewer trips being served in order to avoid paying for more drivers, which leads to lower revenue for the taxi company. For model S2, with bigger fleets, the AT system and the human driver system have the same profit, which indicates that the full-service scheme has more influence than adding human driver costs in the objective function.

The taxi system generates the problem of vehicle imbalance due to the one-way nature of the trips, however, in the case of AVs, it is possible to do the relocation with lower costs between different areas because there is no need for a driver. Service area design and trip selection are able to reduce the negative impact of taxi imbalance in this shared system, therefore, contributing to the maximization of the profit. When the model is free to choose trips (S1), having electric ATs constrains the system in accepting more requests for small fleets because these cars need to charge thus not satisfying trips in the meantime.

It should be noted that the charging activities were simplified in the model formulation. We considered the electric power consumption as a linear function of the AT's driving time and the charged energy also fully proportional to the charging time. However, we do recognize that this is different in reality, which may influence the applicability of this model. Moreover, we assumed that the ATs can directly use the existing road network without any hindrances.

This is a limitation of this part of research since the current road network is not designed for automated driving but human drivers. The infrastructure needs upgrades or adjustments to enable ATs to travel along the road (Lu, 2018; Madadi et al., 2018).

Research question 2: How to choose the ATs' routes considering traffic congestion?

To answer this research question, we proposed an IP model to study how traffic congestion could lead to different route choices of ATs. In the previous study, the travel times between origin and destinations were assumed to be static while varying along the day. Here we considered the influence of traffic congestion by involving flow dependent travel times. The AT system is optimized with the objective to maximize the total profit of such system by deciding on each AT's routing selection. The model was applied to two case studies: a small one with a network of 9 nodes and 12 links for 10 time-step service period and 20 travel requests; and a real network of the city of Delft, the Netherlands, with 46 nodes and 66 links for 15 hours service period and 22,240 travel requests.

Knowing that real-life travel conditions in urban networks show dynamic travel time due to traffic congestion, application of a static minimum travel time vehicle routing model will overestimate the system profitability up to 56%. The influence of considering dynamic travel times in the ATs routing can be seen from two perspectives: the time delay on the congested road links and the extra travel distance caused by the detouring of the vehicles. Comparing the static travel time scenarios with dynamic ones, it is easy to find that the AT system is always more profitable when considering static travel times. The optimization results showed that the AT passengers will experience an arrival delay when considering the traffic congestion generated by the ATs themselves. This is because when the travel time is always the minimum value, the system would take less time than the dynamic travel time scenario, which makes it possible for the ATs to serve more trips and get more revenues from those trips. However, it cannot reflect the real traffic status and the impact of ATs on travel time.

The limitation of this part of research is that we did not consider background traffic flow since we assumed ATs would replace all modes of personal transport and alternative transport modes e.g. metro, bus and tram are usually seen as not contributing to the congestion in the network. When ATs are used in a mixed traffic scenario, then the model will be more complex.

Research question 3: How to satisfy people's real-time demand during the day with an AT system?

To answer this research question, we established a rolling-horizon framework to involve real-time requests and optimize vehicle routing horizon-by-horizon. In the previous study, the demand was assumed to be pre-known, while on this one the ATs were considered to be able to serve not only the pre-booked demand but also the real-time demand. In each horizon, we considered dynamic travel times and performed system optimization with the objective of maximizing the daily profit of such a system through deciding on each AT route. The model

was then applied to a case study in the city of Delft, the Netherlands, with a 15 hours service period and 22,240 travel requests on a road network with 46 nodes and 66 links.

Rolling-horizon was used as a framework for adapting the supply to the real-time requests that keep popping up throughout the day. We divided an operation day into 60 horizons, each of which contains 30 min, and we set the rolling length as 15 min. In practice, this means that when the system makes routing decisions for the ATs, it looks forward as far as 30 min and updates the ATs' status every 15 min. Knowing that travel demand and vehicle status may change in real-time, using a rolling-horizon framework can simulate the real-life situation in a realistic way.

The time distribution of the urban travel demand is not uniformly distributed during the day, which creates the "peak hours" and the "off-peak hours" in the rolling horizon framework. When 600 ATs are applied to the city of Delft, in the morning peak hours, the AT system has a relatively low satisfaction rate. The afternoon peak hours have a similar tendency. However, these ATs provide a higher share of ATs for the low-demand time periods and some of them have a 100% satisfaction rate. This means that with 600 ATs, the high demand for the morning and afternoon peak hours exceeds the transport capacity of the system, while in the off-peak horizons the transport capacity is redundant.

Research question 4: How to solve the route choice problem with ATs in an efficient way?

To answer this research question, we proposed two methods to solve the ATs' route choice model. The model considers the demand as real-time requests which have to be managed via a rolling-horizon framework as explained before. Even though the rolling-horizon framework makes it possible to decrease the scale of the problem in each optimization process, the proposed model still involves a set of non-linear constraints having a flow based travel time function, which makes it an NP-hard one. To tackle this, first, we linearized the non-linear travel time function by transforming it into a set of discrete points and solved it by a commercial solver. In another method, we kept the non-linear constraints and proposed a customized Lagrangian relaxation algorithm to solve it. In this way, it is not only possible to find a near optimal solution but also to have a notion of how good this solution is. The model and the solving algorithm was applied to the case study city of Delft, The Netherlands, with 15 hours service period and 22,240 travel requests on a road network with 46 nodes and 66 links.

The linearization method is an effective way to make the model solvable by commercial software when dealing with a small case. However, this sacrifices the model's accuracy to some extent, since the values of traffic flow and travel time between adjacent break-points cannot be considered. So we proposed another method which can solve the AT's route choice with keeping the non-linear travel time function. The customized Lagrangian solving algorithm is able to solve the proposed NP-hard problem and obtain a good quality feasible solution within an acceptable computation time. Nevertheless, the solving times are still too

high to allow the application of the model in real systems but we believe that this is an initial version of the type of algorithms that must be built and run in the future.

6.2 Main contributions

Based on the findings summarized in section 6.1, we are able to identify the main contributions of this thesis.

The optimization based framework of ATs' planning and operation fills in the theoretical gap in the field of AV application in public transport.

This thesis provides an analytical framework, including a series of mathematical models, to optimize the service area and the route choice of an AT system. This framework can help researchers, industries and policy-makers in two perspectives: firstly, it can be used to test an AT system's performance such as service coverage and operational costs; secondly, this framework could reveal the AT system's potential impacts on the urban road network, e.g., congestions and travel time.

The optimization results depict the vision for how ATs could contribute to the urban transport system in the future.

We applied the proposed models to the case study city of Delft, the Netherlands, and obtained the optimal results for the system operation. This provides an insight into the quantitative benefits for relevant stakeholders when implementing AVs as taxis in urban mobility. The profit of running an AT system, taking into account investment, operation costs, and taxi fees is maximized and analysed in different scenarios. The results gained from the optimization models might not be realistic and actually never be able to be achieved in a real-world application, but they allow the following researchers to use the results to understand what is the best scenario of the AT system regarding the service performance and profitability.

6.3 Implications for practice

This thesis has several important implications and contributions which are useful for the adoption of a fleet of ATs in public transport in the future. The findings may interest researchers, industries and policy-makers in developing an AT system to optimally service people's travel demand. In this section, we discuss the practical implications for future implementation of ATs.

If the parking and charging facilities are sufficient at the train station, using a fleet of 60 ATs could satisfy almost 100% of the travel demand for last mile in Delft Zuid station and bring around 3000 euro as the daily profit with the free-service scheme, while a fleet of 80 ATs is

needed to satisfy a similar number of requests and generate 2691 euro with the full-service scheme.

When the system is free to choose requests to serve (S1), the daily profit is always higher than the scheme with full service (S2), because the latter cannot reject inconvenient trips for the system without cancelling a whole operational zone. S1 is flexible and profitable but will lead to unhappy customers because they may be in a situation in which they have their trips rejected but they know that some ATs are usually available nearby. However model S2 has the advantage of fully covering a zone hence it guarantees the level of service at the expense of a lower profit. Compared to S1, this scheme provides a favourable taxi service which assures that no requests will be missed from the served zones.

When using 600 ATs to provide service in the city of Delft and allowing ride-sharing between passengers who have aligned origins and destinations, this AT system is able to satisfy about 65% of the total travel demand and make a profit of 90.4×103 euro in a day, while each request will have 1.4 minute delay on average due to traffic congestion.

With ride-sharing, the AT system has more capacity to provide better service regarding the number of satisfied requests and the average delay time. Due to the great number of travel requests in Delft especially generated in morning and afternoon peak hours, the passengers who get rejected by the AT system need to use alternative travel modes e.g. by tram, by bus and by bike to accomplish their trips. In fact, the uneven time distribution of the demand and the strict departure and arrival time windows are all important factors in influencing the AT share.

6.4 Recommendations for future research

In this final section, we aim to provide some possible directions for future research.

The AT system usually involves routing vehicles in a network, which consists of a large number of nodes and links, to satisfy real-time transport demands. Therefore, efficient frameworks and algorithms are needed to provide good solutions within an acceptable computation time. This thesis applied the rolling horizon approach to decompose the problem in the time dimension. Other decomposition approaches, e.g. decompose the problem geographically; and other modelling techniques, e.g., multi-stage modelling, could be developed and tested to reduce the scale of the problem. Meanwhile, heuristic algorithms are promising to solve AT systems' scheduling. The AT system requires a timely reaction to the customers' demand, therefore, heuristics with a short computation time and high-quality solutions will considerably improve the ATs' dispatching.

A simulation model should be developed for the AT system to evaluate the results from the optimization model. Compared with optimization methods, the simulation could reflect the dynamics in the system such as the demand, travel time, vehicle's faults, etc. Moreover, the

simulation system could visually demonstrate the performance of the AT system with different fleets, traffic conditions, service schemes, dispatching rules, etc. Therefore, a comprehensive simulation system could be used to compare with the results gained from the optimization model, which aims to test the reliability and viability of the chosen model and perform uncertainty and randomness of this system.

It could be relevant to test the AT system in a larger city with a more complex road network and more public transport demand. This thesis mainly focused on implementing the AT system in Delft with tens of service zones and approximately 20 thousand requests per day. If the system becomes larger, it may lead to difficulties in dispatching the fleet to satisfy the demand. Nevertheless, at the same time, the operation costs could be reduced given the economies of scale.

There is also the possibility of involving choice modelling to analyse people's preferences for using ATs. In this thesis, we treat the demand of using ATs as fixed, indicating that the influence of time delay on travel demand is neglected. In future modelling, ATs cannot be treated as an isolated transport mode anymore. They should be regarded as a part of the urban transport system. The interactions between ATs and other travel modes like public transport, bikes, and walk should be taken into account. Therefore, a shared-use vehicle routing model should be established incorporating mode choice that will decide which clients prefer to take an AT or another competing mode. This model should be able to consider the impact of traffic congestion on the mode split. It assumes that the demand for ATs is elastic depending on its service level considering the competition between ATs and other modes like bikes or walk. The final objective is to achieve an equilibrium between these modes when assigning all passengers to finish their trips.

The operation structure of the AT system should be further studied. This thesis proposed a centralized dispatching system, in which a single operator collects the demand information and schedules the routing for all the vehicles aiming at a global optimum. Meanwhile, a distributed dispatching system, which consists of multiple taxi companies or dispatchers in the same area, should be investigated as the competition and coordination between different companies may affect the performance of the AT system. Therefore, future research should explore operation structures and indicate how to improve system performance in different cases.

Bibliography

Agatz, N., Erera, A., Savelsbergh, M., Wang, X., 2012. Optimization for dynamic ride-sharing: A review. European Journal of Operational Research 223, 295–303. doi:10.1016/j.ejor.2012.05.028

Agatz, N.A.H., Erera, A.L., Savelsbergh, M.W.P., Wang, X., 2011. Dynamic ride-sharing: A simulation study in metro Atlanta. Transportation Research Part B: Methodological 45, 1450–1464. doi:10.1016/j.trb.2011.05.017

An, K., Lo, H.K., 2015. Robust transit network design with stochastic demand considering development density. Transportation Research Part B: Methodological 81, 737–754. doi:10.1016/j.trb.2015.05.019

An, K., Xie, S., Ouyang, Y., 2017. Reliable sensor location for object positioning and surveillance via trilateration. Transportation Research Part B: Methodological 23, 228–245. doi:10.1016/j.trb.2017.11.012

Anderson, J., Kalra, N., Stanley, K., Sorensen, P., Samaras, C., Oluwatola, O., 2016. Autonomous Vehicle Technology: A Guide for Policymakers. doi:10.7249/RR443-2

Arem, B. Van, Driel, C.J.G. Van, Visser, R., 2006. The Impact of Cooperative Adaptive Cruise Control on Traffic-Flow Characteristics. IEEE Transactions on Intelligent Transportation Systems 7, 429–436.

Atasoy, B., Ikeda, T., Ben-Akiva, M.E., 2015. Optimizing a Flexible Mobility on Demand System. Transportation Research Record: Journal of the Transportation Research Board 2536, 76–85. doi:10.3141/2536-10

Bai, Y., Hwang, T., Kang, S., Ouyang, Y., 2011. Biofuel refinery location and supply chain planning under traffic congestion. Transportation Research Part B: Methodological 45, 162–175. doi:10.1016/j.trb.2010.04.006

Barter, P., 2013. Cars are parked 95% of the time [WWW Document]. Reinventingparking.Org. URL https://www.reinventingparking.org/2013/02/cars-are-parked-95-of-time-lets-check.html

Bates, J., Leibling, D., 2012. Spaced Out Perspectives on parking policy, RAC Foundation. IEEE. doi:10.1109/CMD.2008.4580239

Beckmann, M.J., McGuire, C.B., Winsten, C.B., 1955. Studies in the Economics of Transportation. Yale University Press, New Haven. doi:10.1057/jors.1980.83

Bierstadt, J., Gooze, A., Gray, C., Peterman, J., Raykin, L., Walters, J., 2014. Effects of Next - Generation Vehicles on Travel Demand and Highway Capacity.

Bose, A., Ioannou, P. a., 2003. Analysis of traffic flow with mixed manual and semiautomated vehicles. IEEE Transactions on Intelligent Transportation Systems 4, 173–188. doi:10.1109/TITS.2003.821340

Braekers, K., Caris, A., Janssens, G.K., 2014. Exact and meta-heuristic approach for a general heterogeneous dial-a-ride problem with multiple depots. Transportation Research Part B: Methodological 67, 166–186. doi:10.1016/j.trb.2014.05.007

Braekers, K., Ramaekers, K., Van Nieuwenhuyse, I., 2016. The vehicle routing problem: State of the art classification and review. Computers and Industrial Engineering 99, 300–313. doi:10.1016/j.cie.2015.12.007

Brannlund, U., Lindberg, P.O., Nou, a., Nilsson, J.-E., 1998. Railway Timetabling Using Lagrangian Relaxation. Transportation Science 32, 358–369. doi:10.1287/trsc.32.4.358

Calvert, S.C., Van Den Broek, T.H.A., Van Noort, M., 2011. Modelling cooperative driving in congestion shockwaves on a freeway network. IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC 614–619. doi:10.1109/ITSC.2011.6082837

Celsor, C., Millard-Ball, A., 2007. Where Does Carsharing Work?: Using Geographic Information Systems to Assess Market Potential. Transportation Research Record: Journal of the Transportation Research Board 1992, 61–69. doi:10.3141/1992-08

Chong, Z.J., Qin, B., Bandyopadhyay, T., Wongpiromsarn, T., Rankin, E.S., Ang, M.H., Frazzoli, E., Rus, D., Hsu, D., Low, K.H., 2011. Autonomous personal vehicle for the first- and last-mile transportation services. Proceedings of the 2011 IEEE 5th International Conference on Cybernetics and Intelligent Systems, CIS 2011 253–260. doi:10.1109/ICCIS.2011.6070337

Cordeau, J.F., Laporte, G., 2007. The dial-a-ride problem: Models and algorithms. Annals of Operations Research 153, 29–46. doi:10.1007/s10479-007-0170-8

Correia, G., Viegas, J.M., 2011. Carpooling and carpool clubs: Clarifying concepts and assessing value enhancement possibilities through a Stated Preference web survey in Lisbon, Portugal. Transportation Research Part A: Policy and Practice 45, 81–90. doi:10.1016/j.tra.2010.11.001

Correia, G.H. de A., Looff, E., van Cranenburgh, S., Snelder, M., van Arem, B., 2019. On the impact of vehicle automation on the value of travel time while performing work and leisure activities in a car: Theoretical insights and results from a stated preference survey. Transportation Research Part A: Policy and Practice 119, 359–382.

doi:10.1016/j.tra.2018.11.016

Correia, G.H. de A., Santos, R.F.G., 2014. Optimizing the Use of Electric Vehicles in a Regional Car Rental Fleet. Transportation Research Record: Journal of the Transportation Research Board 2454, 76–83. doi:10.3141/2454-10

Correia, G.H. de A., van Arem, B., 2016a. Solving the User Optimum Privately Owned Automated Vehicles Assignment Problem (UO-POAVAP): A model to explore the impacts of self-driving vehicles on urban mobility. Transportation Research Part B: Methodological 87, 64–88. doi:10.1016/j.trb.2016.03.002

Correia, G.H. de A., van Arem, B., 2016b. Trips and network of the case-study city of Delft. doi:10.13140/RG.2.2.11097.83043

Correia, G.H.D.A., Antunes, A.P., 2012. Optimization approach to depot location and trip selection in one-way carsharing systems. Transportation Research Part E: Logistics and Transportation Review 48, 233–247. doi:10.1016/j.tre.2011.06.003

Cui, T., Ouyang, Y., Shen, Z.-J.M., 2010. Reliable Facility Location Design Under the Risk of Disruptions. Operations Research 58, 998–1011. doi:10.1287/opre.1090.0801

Dafermos, S., Sparrow, F., 1969. Traffic assignment problem for a general network. U S Bur Standards-J Research-Mathematical Sciences.

Daimler, A., 2017. Mercedes-Benz Trucks: Safety: New assistance systems: Active Brake Assist 4 emergency: braking assistant featuring pedestrian recognition and: Sideguard Assist [WWW Document]. Daimler. URL https://media.daimler.com/marsMediaSite/ko/en/12367326

Dios Ortúzar, J. de., Willumsen, L.G., 2000. Modelling transport. Wiley, Chichester [u.a.].

Discover Renault Twizy [WWW Document], 2014. URL http://www.renault.fr/gamme-renault/vehicules-electriques/twizy/twizy/

Fagnant, D.J., Kockelman, K., 2013. Preparing a Nation for Autonomous Vehicles: Opportunities, Barriers and Policy Recommendations for Capitalizing on Self-Driven Vehicles. Transportation Research Part A: Policy and Practice 77, 167–181. doi:10.1016/j.tra.2015.04.003

Fagnant, D.J., Kockelman, K.M., 2014. The travel and environmental implications of shared autonomous vehicles, using agent-based model scenarios. Transportation Research Part C: Emerging Technologies 40, 1–13. doi:10.1016/j.trc.2013.12.001

FICO, 2014. Getting Started with Xpress Release 7.7.

Figliozzi, M., 2011. Vehicle Routing Problem for Emissions Minimization. Transportation Research Record: Journal of the Transportation Research Board 2197, 1–7. doi:10.3141/2197-01

Fisher, M.L., 2004. The Lagrangian Relaxation Method for Solving Integer Programming Problems. Management Science 50, 1861–1871. doi:10.1287/mnsc.1040.0263

Fisher, M.L., 1981. The Lagrangian Relaxation Method for Solving Integer Programming Problems. Management Science 27, 1–18. doi:10.1287/mnsc.27.1.1

Ford, H.J., 2012. Shared Autonomous Taxis : Implementing an Efficient Alternative to Automobile Dependency. Princenton University.

Furuhata, M., Dessouky, M., Ordóñez, F., Brunet, M.E., Wang, X., Koenig, S., 2013. Ridesharing: The state-of-the-art and future directions. Transportation Research Part B: Methodological 57, 28–46. doi:10.1016/j.trb.2013.08.012

Garcia-Najera, a, Bullinaria, J. a, 2010. Bi-objective optimization for the Vehicle Routing Problem with Time Windows: Using route similarity to enhance performance. Evolutionary Multi-Criterion Optimization Lecture Notes in Computer Science Volume 5467, 2009 5467 LNCS, 275–289. doi:10.1007/978-3-642-01020-0_24

Golson, J., Bohn, D., 2016. All new Tesla cars now have hardware for 'full self-driving capabilities' [WWW Document]. The Verge. URL http://www.theverge.com/2016/10/19/13340938/tesla-autopilot-update-model-3-elon-musk-update

Hess, D.B., 2009. Access to Public Transit and Its Influence on Ridership for Older Adults in Two U.S. Cities. Journal of Transport and Land Use 2, 3–27. doi:10.5198/jtlu.v2i1.11

Ho, S.C., Szeto, W.Y., Kuo, Y.H., Leung, J.M.Y., Petering, M., Tou, T.W.H., 2018. A survey of dial-a-ride problems: Literature review and recent developments. Transportation Research Part B: Methodological 111, 395–421. doi:10.1016/j.trb.2018.02.001

Hoogendoorn-Lanser, S., van Nes, R., Hoogendoorn, S.P., Bovy, P.H.L., 2006. Home-activity approach to muli-modal travel choice modelling. Transportation Research Record: Journal of the Transportation Research Board 1985, 180–187.

Hoogendoorn, R., van Arerm, B., Hoogendoom, S., 2014. Automated Driving, Traffic Flow Efficiency, and Human Factors. Transportation Research Record: Journal of the Transportation Research Board 2422, 113–120. doi:10.3141/2422-13

Huang, K., Correia, G.H. de A., An, K., 2018. Solving the station-based one-way carsharing network planning problem with relocations and non-linear demand. Transportation Research Part C: Emerging Technologies 90, 1–17. doi:10.1016/j.trc.2018.02.020

Imai, A., Nishimura, E., Current, J., 2007. A Lagrangian relaxation-based heuristic for the vehicle routing with full container load. European Journal of Operational Research 176, 87–105. doi:10.1016/j.ejor.2005.06.044

Jaw, J.J., Odoni, A.R., Psaraftis, H.N., Wilson, N.H.M., 1986. A heuristic algorithm for the multi-vehicle advance request dial-a-ride problem with time windows. Transportation Research Part B: Methodological 20, 243–257. doi:10.1016/0191-2615(86)90020-2

Jorge, D., Barnhart, C., de Almeida Correia, G.H., 2015a. Assessing the viability of enabling a round-trip carsharing system to accept one-way trips: Application to Logan Airport in Boston. Transportation Research Part C: Emerging Technologies 56, 359–372. doi:10.1016/j.trc.2015.04.020

Jorge, D., Correia, G., Barnhart, C., 2012. Testing the Validity of the MIP Approach for Locating Carsharing Stations in One-way Systems. Procedia - Social and Behavioral Sciences 54, 138–148. doi:10.1016/j.sbspro.2012.09.733

Jorge, D., Correia, G.H.A., Barnhart, C., 2014. Comparing optimal relocation operations with

simulated relocation policies in one-way carsharing systems. IEEE Transactions on Intelligent Transportation Systems 15, 1667–1675. doi:10.1109/TITS.2014.2304358

Jorge, D., Molnar, G., de Almeida Correia, G.H., 2015b. Trip pricing of one-way station-based carsharing networks with zone and time of day price variations. Transportation Research Part B: Methodological 1–22. doi:10.1016/j.trb.2015.06.003

Juster, R., Schonfeld, P., 2013. Comparative Analysis of Personal Rapid Transit as an Urban Transportation Mode. Transportation Research Record: Journal of the Transportation Research Board 2350, 128–135. doi:10.3141/2350-15

Kaufman, D.E., Nonis, J., Smith, R.L., 1992. A mixed integer linear programming formulation of the dynamic traffic assignment problem. Systems, Man and Cybernetics, 1992., IEEE International Conference on 232–235 vol.1. doi:10.1109/ICSMC.1992.271771

KPMG, 2012. Self-driving cars : the next revolution.

Krueger, R., Rashidi, T.H., Rose, J.M., 2016. Preferences for shared autonomous vehicles. Transportation Research Part C: Emerging Technologies 69, 343–355. doi:10.1016/j.trc.2016.06.015

Lamotte, R., de Palma, A., Geroliminis, N., 2017. On the use of reservation-based autonomous vehicles for demand management. Transportation Research Part B: Methodological 99, 205–227. doi:10.1016/j.trb.2017.01.003

Laporte, G., 2009. Fifty Years of Vehicle Routing. Transportation Science 43, 408–416. doi:10.1287/trsc.1090.0301

Lei, C., Ouyang, Y., 2018. Continuous approximation for demand balancing in solving large-scale one-commodity pickup and delivery problems. Transportation Research Part B 109, 90–109. doi:10.1016/j.trb.2018.01.009

Levin, M.W., 2017. Congestion-aware system optimal route choice for shared autonomous vehicles. Transportation Research Part C: Emerging Technologies 82, 229–247. doi:10.1016/j.trc.2017.06.020

Li, B., Krushinsky, D., Van Woensel, T., Reijers, H.A., 2016. The Share-a-Ride problem with stochastic travel times and stochastic delivery locations. Transportation Research Part C: Emerging Technologies 67, 95–108. doi:10.1016/j.trc.2016.01.014

Li, X., Ma, J., Cui, J., Ghiasi, A., Zhou, F., 2016. Design framework of large-scale one-way electric vehicle sharing systems: A continuum approximation model. Transportation Research Part B: Methodological 88, 21–45. doi:10.1016/j.trb.2016.01.014

Liang, X., Correia, G.H. de A., van Arem, B., 2016. Optimizing the service area and trip selection of an electric automated taxi system used for the last mile of train trips. Transportation Research Part E: Logistics and Transportation Review 93, 115–129. doi:10.1016/j.tre.2016.05.006

Light, D., 2012. A Scenario: The End of Auto Insurance [WWW Document]. URL http://www.celent.com/reports/scenario-end-auto-insurance

Litman, T., 2018. Autonomous vehicle implementation predictions : Implications for

Transport Planning.

Long, J., Tan, W., Szeto, W.Y., Li, Y., 2018. Ride-sharing with travel time uncertainty. Transportation Research Part B 118, 143–171. doi:10.1016/j.trb.2018.10.004

Lu, X., 2018. Infrastructure Requirements for Automated Driving. Delft University of Technology.

Luo, Y., Schonfeld, P., 2011. Online Rejected-Reinsertion Heuristics for Dynamic Multivehicle Dial-a-Ride Problem. Transportation Research Record: Journal of the Transportation Research Board 2218, 59–67. doi:10.3141/2218-07

Ma, J., Li, X., Zhou, F., Hao, W., 2017. Designing optimal autonomous vehicle sharing and reservation systems: A linear programming approach. Transportation Research Part C: Emerging Technologies 84, 124–141. doi:10.1016/j.trc.2017.08.022

Madadi, B., Nes, R. Van, Snelder, M., Arem, B. Van, 2019. Case Studies on Transport Policy Assessing the travel impacts of subnetworks for automated driving : An exploratory study. Case Studies on Transport Policy 7, 48–56. doi:10.1016/j.cstp.2018.11.006

Madadi, B., van Nes, R., Snelder, M., van Arem, B., 2018. Image-based assessment of road network readiness for automated driving :

Mahmoudi, M., Zhou, X., 2016. Finding optimal solutions for vehicle routing problem with pickup and delivery services with time windows: A dynamic programming approach based on state-space-time network representations. Transportation Research Part B: Methodological 89, 19–42. doi:10.1016/j.trb.2016.03.009

MARSHALL, A., 2018. Toyota Joins Uber on Its Tortuous Journey to Self-Driving Cars [WWW Document]. URL https://www.wired.com/story/toyota-uber-tortuous-journey-self-driving-cars/

Martinez, L.M., Viegas, J.M., 2017. Assessing the impacts of deploying a shared self-driving urban mobility system: An agent-based model applied to the city of Lisbon, Portugal. International Journal of Transportation Science and Technology 6, 1–15. doi:10.1016/j.ijtst.2017.05.005

Masoud, N., Jayakrishnan, R., 2017. A decomposition algorithm to solve the multi-hop Peer-to-Peer ride-matching problem. Transportation Research Part B: Methodological 99, 1–29. doi:10.1016/j.trb.2017.01.004

Mercedes Blog-Team, 2017. How to Use DISTRONIC PLUS Cruise Control in 2016 Mercedes-Benz [WWW Document]. Mercedes-Benz official blog. URL http://www.mbscottsdale.com/blog/how-to-use-distronic-plus-cruise-control-in-2016-mercedes-benz/

Najmi, A., Rey, D., Rashidi, T.H., 2017. Novel dynamic formulations for real-time ride-sharing systems. Transportation Research Part E: Logistics and Transportation Review 108, 122–140. doi:10.1016/j.tre.2017.10.009

Nieuwenhuijsen, J., Correia, G.H. de A., Milakis, D., van Arem, B., van Daalen, E., 2018. Towards a quantitative method to analyze the long-term innovation diffusion of automated vehicles technology using system dynamics. Transportation Research Part C: Emerging Technologies 86, 300–327. doi:10.1016/j.trc.2017.11.016

NS, 2014. Sprinter SLT [WWW Document]. URL http://www.ns.nl/en/about-ns/trains-of-ns/sprinter-slt.html

Park, S., Lee, T.E., Sung, C.S., 2010. A three-level supply chain network design model with risk-pooling and lead times. Transportation Research Part E: Logistics and Transportation Review 46, 563–581. doi:10.1016/j.tre.2009.12.004

Peeta, S., Mahmassani, H.S., 1995. System optimal and user equilibrium time-dependent traffic assignment in congested networks. Annals of Operations Research 60, 81–113. doi:10.1007/BF02031941

Pillac, V., Gendreau, M., Guéret, C., Medaglia, A.L., 2013. A review of dynamic vehicle routing problems. European Journal of Operational Research 225, 1–11. doi:10.1016/j.ejor.2012.08.015

Psaraftis, H., 1980. A Dynamic Programming solution to the single vehicle many-to-many immediate request dial-a-ride problem. Transportation Science 14, 130–154. doi:10.1287/trsc.14.2.130

Psaraftis, H.N., 1988. Vehicle routing: Methods and studies. Dynamic Vehicle Routing Problems. North Holland, Amsterdam, The Netherlands 223–248.

Pyper, J., 2015. Self-Driving Cars Could Cut Greenhouse Gas Pollution [WWW Document]. Scientificamerican.Com. URL http://www.scientificamerican.com/article/self-driving-cars-could-cut-greenhouse-gas-pollution/

SAE International, 2018. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles.

Salanova, J.M., Estrada, M., Aifadopoulou, G., Mitsakis, E., 2011. A review of the modeling of taxi services. Procedia - Social and Behavioral Sciences 20, 150–161. doi:10.1016/j.sbspro.2011.08.020

Schaller, B., 2018. The New Automobility: Lyft, Uber and the Future of American Cities.

Schuster, T., Byrne, J., Corbett, J., Schreuder, Y., 2005. Assessing the Potential Extent of Carsharing: A New Method and Its Implications. Transportation Research Record: Journal of the Transportation Research Board 1927, 174–181. doi:10.3141/1927-20

Sethi, S., Sorger, G., Kleiner, A., Nebel, B., Ziparo, V.A., 1991. A mechanism for dynamic ride sharing based on parallel auctions. Annals of Operations Research 29, 266–272. doi:10.1007/BF02283607

Shaheen, S., Sperling, D., Wagner, C., 1999. A Short History of Carsharing in the 90's. Journal of World Transport Policy and Practice 5, 16–37. doi:10.1007/s11116-007-9132-x

Shen, Q., Chu, F., Chen, H., 2011. A Lagrangian relaxation approach for a multi-mode inventory routing problem with transshipment in crude oil transportation. Computers and Chemical Engineering 35, 2113–2123. doi:10.1016/j.compchemeng.2011.01.005

Spieser, K., Treleaven, K., Zhang, R., Frazzoli, E., Morton, D., Pavone, M., 2014. Toward a Systematic Approach to the Design and Evaluation of Automated Mobility-on-Demand Systems: A Case Study in Singapore, in: Road Vehicle Automation. pp. 229–245.

doi:10.1007/978-3-319-05990-7_20

Tanatmis, A., Ruzika, S., Kienle, F., 2010. A Lagrangian relaxation based decoding algorithm for LTE turbo codes. 6th International Symposium on Turbo Codes and Iterative Information Processing, ISTC 2010 369–373. doi:10.1109/ISTC.2010.5613906

TCA, 2018. Driver Shortage in Japan Speeds Up Commercial Use of Self-Driving Tech [WWW Document]. Transport Topics. URL https://www.ttnews.com/articles/driver-shortage-japan-speeds-commercial-use-self-driving-tech

The Tesla Team, 2018. Q3 2018 Vehicle Safety Report [WWW Document]. URL https://www.tesla.com/blog/q3-2018-vehicle-safety-report

Uber [WWW Document], 2017. URL https://www.uber.com/

Wang, U., 2015. Are Self-Driving Vehicles Good for the Environment? [WWW Document]. doi:https://ensia.com/features/are-self-driving-vehicles-good-for-the-environment/

Ward, J., 2014. Mercedes-Benz E400: Distronic Plus with Steering Assist demonstrated [WWW Document]. URL https://www.caradvice.com.au/290478/mercedes-benz-e400-distronic-plus-with-steering-assist-demonstrated/

Wardrop, J.G., 1952. ROAD PAPER. SOME THEORETICAL ASPECTS OF ROAD TRAFFIC RESEARCH. Proceedings of the Institution of Civil Engineers 1, 325–362. doi:10.1680/ipeds.1952.11259

Waymo, 2018. Waymo: What’s Next [WWW Document]. URL https://waymo.com/whats-next/

Yang, J., Jaillet, P., Mahmassani, H., 1999. On-Line Algorithms for Truck Fleet Assignment and Scheduling Under Real-Time Information. Transportation Research Record 1667, 107–113. doi:10.3141/1667-13

Yang, L., Zhou, X., 2014. Constraint reformulation and a Lagrangian relaxation-based solution algorithm for a least expected time path problem. Transportation Research Part B: Methodological 59, 22–44. doi:10.1016/j.trb.2013.10.012

Yap, M.D., Correia, G., van Arem, B., 2016. Preferences of travellers for using automated vehicles as last mile public transport of multimodal train trips. Transportation Research Part A: Policy and Practice 94, 1–16. doi:10.1016/j.tra.2016.09.003

Zhang, B., 2014. Autonomous Cars May Save US A Trillion Dollars [WWW Document]. URL https://www.businessinsider.com/morgan-stanley-autonomous-cars-trillion-dollars-2014-9?international=true&r=US&IR=T

Zhou, W., Teng, H., 2016. Simultaneous passenger train routing and timetabling using an efficient train-based Lagrangian relaxation decomposition. Transportation Research Part B: Methodological 94, 409–439. doi:10.1016/j.trb.2016.10.010

Summary

In recent years, technology development has accelerated the future roll-out of vehicle automation. An automated vehicle (AV), also known as a driverless car and a self-driving car is an advanced type of vehicle that can drive itself on existing roads. A possible area of application for AVs is public transport. The concept of automated taxis (ATs) is supposed to offer a seamless door-to-door service within a city area for all passengers. With automation technology maturing, we may be able to see the situation in which hundreds or even thousands of ATs will be on the road replacing private vehicles accounting for the majority of people's daily trips. However, little attention has been devoted to the usage of a fleet of ATs and their effect on a real-scale road network.

In this thesis, we explore how automated driving can serve mobility and what is the best way to introduce this technology as part of the existing transport networks. This is also the research gap this thesis is going to fill. The objective of this thesis is to contribute to the planning and operational strategies that these AT systems should follow in order to satisfy mobility demand.

This thesis uses mathematical optimization to address the above research problems. A mathematical optimization problem consists of maximizing or minimizing a function by systematically selecting some input values within a defined domain. It aims to find the best available values of the objective function and the corresponding values of the problem input. The purpose of this thesis is to provide a tool to support the decision-making processes both for long-term planning strategies and short-term tactical operations when ATs are going to be applied in the urban transport system.

Firstly, we focus on defining the service area of ATs according to the travel demand distribution in time and space during the planning stage of adopting an AT system. A facility

location problem is formulated as an integer programming (IP) model, to select several geographical zones as the service area of an AT system for the last mile of train trips. Based on the results, we are able to conclude that service area design and trip selection are able to reduce the negative impact of taxi imbalance in this shared system, therefore, contributing to the maximization of the profit.

Secondly, we focus on optimizing the route choice when ATs are assigned to pick-up and deliver passengers during the operation stage. This formulation involves traffic assignment with congestion for the AT fleet to decide the route choice, by integrating a non-linear flow based travel time function. Knowing that real-life travel conditions in urban networks show dynamic travel time due to traffic congestion, application of a static minimum travel time vehicle routing model will overestimate the system profitability.

Thirdly, we focus on establishing a framework to address the real-time demand of ATs. This makes ATs' route choice results flexible to handle the new customers being generated through time. A rolling-horizon framework is proposed to address real-time requests, which is called a dynamic dial-a-ride problem (DARP). Since travel demand and vehicle status may change in real-time, using a rolling-horizon framework can simulate the real-life situation in a realistic way.

Fourthly, we focus on the mathematical challenge of solving the problems proposed in this thesis. We proposed two methods to solve the ATs' route choice model: 1) linearizing the non-linear travel time function and solving it by a commercial solver; 2) keeping the non-linear constraints and proposed a customized Lagrangian relaxation algorithm to solve it. The linearization method is an effective way to make the model solvable by commercial software when dealing with a small case, while a customized Lagrangian relaxation solution algorithm is possible to find a near optimal solution when solving a large scale problem.

To sum up, this thesis provides an analytical framework, including a series of mathematical models, to optimize the service area and the route choice of an AT system. We applied the proposed models to the case study city of Delft, the Netherlands, and obtained the optimal results for the system operation. This provides an insight into the quantitative benefits for relevant stakeholders when implementing ATs in urban mobility.

Samenvatting

De laatste jaren hebben ontwikkelingen in de technologie het op de markt brengen van geautomatiseerd rijden versneld. Een automatisch voertuig (AV), oftewel de autonome of de zelfrijdende auto, is een geavanceerd voertuig dat zelfstandig op bestaande wegen kan rijden. Een mogelijk toepassingsgebied voor AV's is openbaar vervoer. Het concept van geautomatiseerde taxi's (AT's) zou alle passagiers binnen een stadsgebied een service van deur tot deur moeten bieden. Nu de automatiseringstechnologie steeds beter wordt, kan het zijn dat we het straks meemaken dat honderden of zelfs duizenden AT's op de openbare weg rondrijden en de privévoertuigen vervangen waarmee mensen nu het merendeel van hun dagelijkse ritten maken. Er is echter nog weinig aandacht besteed aan het gebruik van een wagenpark ATs en welk effect die op een echt wegennet zouden hebben.

In dit proefschrift onderzoeken we welke rol automatisch rijden kan spelen in de mobiliteit en wat de beste manier is om deze technologie binnen de bestaande transportnetwerken te introduceren. Dat is ook de leemte in het onderzoek dat dit proefschrift zal opvullen. Dit proefschrift beoogt bij te dragen aan de strategieën voor de planning en uitvoering die deze AT-systemen moeten volgen om in de mobiliteitsvraag te voorzien.

In dit proefschrift wordt wiskundige optimalisatie gebruikt om bovenstaande onderzoeksvragen te beantwoorden. Een wiskundig optimalisatieprobleem bestaat uit het maximaliseren of minimaliseren van een functie door systematisch inputwaarden uit een gedefinieerd domein te variëren. Er wordt getracht de best beschikbare waarden van de doelfunctie te vinden, evenals de overeenkomstige waarden van de input van het probleem. Dit proefschrift beoogt te voorzien in een instrument dat de besluitvormingsprocessen

ondersteunt voor zowel strategieën voor langetermijnplanning als voor de bedrijfsfase op de korte termijn wanneer AT's in het stedelijke transportsysteem worden ingezet.

Ten eerste concentreren we ons op het definiëren van het verzorgingsgebied van AT's op basis van de distributie in tijd en ruimte van de vraag naar vervoer tijdens de planningsfase van het invoeren van een AT-systeem. Een faciliteitenlocatieprobleem wordt gedefinieerd als een integer programmeringsmodel (IP-model) om verschillende geografische zones als het verzorgingsgebied van een AT-systeem te kiezen voor de laatste anderhalve kilometer van treinreizen. Op basis van de resultaten kunnen we concluderen dat het ontwerp van het verzorgingsgebied en de reiskeuze de negatieve impact van een onbalans van taxi's in dit gedeelde systeem kunnen verminderen, wat bijdraagt aan maximalisering van de winst.

Ten tweede concentreren we ons op optimalisatie van de routekeuze als AT's passagiers moeten ophalen en afleveren tijdens de bedrijfsfase. Deze formule omvat het inzetten van het AT-wagenpark bij opstoppingen om de routekeuze te bepalen door een niet-lineaire, op de doorstroming gebaseerde reistijdfunctie te integreren. Aangezien de werkelijke reisomstandigheden in stedelijke netwerken een dynamische reistijd vertonen vanwege opstoppingen, zal de winstgevendheid van het systeem worden overschat als er een routemodel met een statische minimale reistijd wordt toegepast.

Ten derde richten we ons op het bepalen van een tijdsdimensiekader om de werkelijke vraag naar AT's te bepalen. Hierdoor worden de resultaten van de door de AT gekozen route flexibel genoeg om nieuwe klanten te kunnen inpassen die door de tijd heen gegenereerd worden. We stellen een zogenaamd 'rolling horizon'-kader voor om aanvragen onmiddellijk af te handelen; het 'dynamic dial-a-ride problem' (DARP). Aangezien de vraag naar vervoer en de status van voertuigen op het moment zelf kan veranderen, kan het gebruik van een dergelijk 'rolling horizon'-kader de werkelijke situatie realistisch nabootsen.

Ten vierde richten we ons op de wiskundige uitdaging van het oplossen van de problemen die in dit proefschrift worden genoemd. We stellen twee methoden voor als oplossing voor het routekeuze model van de AT: 1) de non-lineaire reistijdfunctie lineariseren en met commerciële software op te lossen; 2) de non-lineaire beperkingen behouden en een Lagrange relaxatie-algoritme voorstellen om deze op te lossen. De linearisatiemethode is een effectieve manier om het model door commerciële software te kunnen laten oplossen bij een klein probleem, terwijl bij een grootschalig probleem een 'bijna-optimale' oplossing kan worden gevonden met een Lagrange relaxatie-algoritme.

Ter samenvatting: dit proefschrift biedt een analytisch kader, waaronder een aantal wiskundige modellen, om het verzorgingsgebied en de routekeuze van een AT-systeem te optimaliseren. We hebben de voorgestelde modellen als casestudy toegepast op de stad Delft en hebben de optimale resultaten voor de bedrijfsfase van het systeem verkregen. Dit biedt inzicht in de kwantitatieve voordelen voor relevante stakeholders wanneer ATs voor vervoer in de stad worden ingevoerd.

TRAIL Thesis Series

The following list contains the most recent dissertations in the TRAIL Thesis Series. For a complete overview of more than 250 titles see the TRAIL website: www.rsTRAIL.nl.

The TRAIL Thesis Series is a series of the Netherlands TRAIL Research School on transport, infrastructure and logistics.

Liang, X., *Planning and Operation of Automated Taxi Systems*, T2019/13, September 2019, TRAIL Thesis Series, the Netherlands

Ton, D., *Unravelling Mode and Route Choice Behaviour of Active Mode Users*, T2019/12, September 2019, TRAIL Thesis Series, the Netherlands

Shu, Y., *Vessel Route Choice Model and Operational Model Based on Optimal Control*, T2019/11, September 2019, TRAIL Thesis Series, the Netherlands

Luan, X., *Traffic Management Optimization of Railway Networks*, T2019/10, July 2019, TRAIL Thesis Series, the Netherlands

Hu, Q., *Container Transport inside the Port Area and to the Hinterland*, T2019/9, July 2019, TRAIL Thesis Series, the Netherlands

Andani, I.G.A., *Toll Roads in Indonesia: transport system, accessibility, spatial and equity impacts*, T2019/8, June 2019, TRAIL Thesis Series, the Netherlands

Ma, W., *Sustainability of Deep Sea Mining Transport Plans*, T2019/7, June 2019, TRAIL Thesis Series, the Netherlands

Alemi, A., *Railway Wheel Defect Identification*, T2019/6, January 2019, TRAIL Thesis Series, the Netherlands

Liao, F., *Consumers, Business Models and Electric Vehicles*, T2019/5, May 2019, TRAIL Thesis Series, the Netherlands

Tamminga, G., *A Novel Design of the Transport Infrastructure for Traffic Simulation Models*, T2019/4, March 2019, TRAIL Thesis Series, the Netherlands

Lin, X., *Controlled Perishable Goods Logistics: Real-time coordination for fresher products*, T2019/3, January 2019, TRAIL Thesis Series, the Netherlands

Dafnomilis, I., *Green Bulk Terminals: A strategic level approach to solid biomass terminal design*, T2019/2, January 2019, TRAIL Thesis Series, the Netherlands

Feng, Fan, *Information Integration and Intelligent Control of Port Logistics System*, T2019/1, January 2019, TRAIL Thesis Series, the Netherlands

Beinum, A.S. van, *Turbulence in Traffic at Motorway Ramps and its Impact on Traffic Operations and Safety,* T2018/12, December 2018, TRAIL Thesis Series, the Netherlands

Bellsolà Olba, X., *Assessment of Capacity and Risk: A Framework for Vessel Traffic in Ports,* T2018/11, December 2018, TRAIL Thesis Series, the Netherlands

Knapper, A.S., *The Effects of using Mobile Phones and Navigation Systems during Driving,* T2018/10, December 2018, TRAIL Thesis Series, the Netherlands

Varotto, S.F., *Driver Behaviour during Control Transitions between Adaptive Cruise Control and Manual Driving: empirics and models,* T2018/9, December 2018, TRAIL Thesis Series, the Netherlands

Stelling-Kończak, A., *Cycling Safe and Sound,* T2018/8, November 2018, TRAIL Thesis Series, the Netherlands

Essen, van M.A., *The Potential of Social Routing Advice,* T2018/7, October 2018, TRAIL Thesis Series, the Netherlands

Su, Zhou, *Maintenance Optimization for Railway Infrastructure Networks*, T2018/6, September 2018, TRAIL Thesis Series, the Netherlands

Cai, J., *Residual Ultimate Strength of Seamless Metallic Pipelines with Structural Damage,* T2018/5, September 2018, TRAIL Thesis Series, the Netherlands

Ghaemi, N., *Short-turning Trains during Full Blockages in Railway Disruption Management,* T2018/4, July 2018, TRAIL Thesis Series, the Netherlands

Gun, van der J.P.T., *Multimodal Transportation Simulation for Emergencies using the Link Transmission Model,* T2018/3, May 2018, TRAIL Thesis Series, the Netherlands

Van Riessen, B., *Optimal Transportation Plans and Portfolios for Synchromodal Container Networks,* T2018/2, March 2018, TRAIL Thesis Series, the Netherlands

Saeedi, H., *Network-Level Analysis of the Market and Performance of Intermodal Freight Transport,* T2018/1, March 2018, TRAIL Thesis Series, the Netherlands

Ypsilantis, P., *The Design, Planning and Execution of Sustainable Intermodal Port-hinterland Transport Networks,* T2017/14, December 2017, TRAIL Thesis Series, the Netherlands

Han, Y, *Fast Model Predictive Control Approaches for Road Traffic Control,* T2017/13, December 2017, TRAIL Thesis Series, the Netherlands

Wang, P., *Train Trajectory Optimization Methods for Energy-Efficient Railway Operations,* T2017/12, December 2017, TRAIL Thesis Series, the Netherlands

Weg, G.S. van de, *Efficient Algorithms for Network-wide Road Traffic Control,* T2017/11, October 2017, TRAIL Thesis Series, the Netherlands

He, D., *Energy Saving for Belt Conveyors by Speed Control,* T2017/10, July 2017, TRAIL Thesis Series, the Netherlands

Curriculum Vitae

Xiao Liang was born on July 27, 1988 in Liaoning, China. She studied Transportation Engineering at Southwest Jiaotong University in Chengdu, China and obtained her B.Sc degree in 2011. After that, she obtained her M.Sc degree in Communication and Transportation Engineering in 2014 from Tongji University in Shanghai, China.

In September 2014, she started her PhD at Department Transport & Planning, Delft University of Technology in Delft, the Netherlands. Her PhD project focuses on the optimization of an automated taxi system in planning and operation level. In February 2018, she visited Monash University in Australia for two months, working on the collaboration research about elastic demand of automated taxis.

In March 2019, she started working as a postdoc researcher at Department Transport & Planning, Delft University of Technology in Delft, the Netherlands. She works on the Research and Innovation (R&I) project “ELECTRIC TRAVELLING” which intends to ease the implementation and further development of electromobility (e-mobility) in urban and suburban areas.

Her research interests include operations research, optimization, transport modelling, vehicle routing, choice modelling.

Publications

Journal papers

1. **X. Liang**, G. H. de A. Correia, and B. van Arem. Applying a Model for Trip Assignment and Dynamic Routing of Automated Taxis with Congestion: System Performance in the City of Delft, The Netherlands. Transportation Research Record: Journal of the Transportation Research Board, May 2018, p. 036119811875804.

2. **X. Liang**, G. H. de A. Correia, and B. Van Arem. An optimization model for vehicle routing of automated taxi trips with dynamic travel times. Transportation Research Procedia, Vol. 27, 2017, pp. 736–743.

3. **X. Liang**, G. H. de A. Correia, and B. van Arem. Optimizing the service area and trip selection of an electric automated taxi system used for the last mile of train trips. Transportation Research Part E: Logistics and Transportation Review, Vol. 93, Sep. 2016, pp. 115–129.

4. **X. Liang**, G. H. de A. Correia, K. An, and B. van Arem. Automated taxis' dial-a-ride problem with ride-sharing considering dynamic travel times: an application to Delft, The Netherlands. (under review)

5. **X. Liang** and G. H. de A. Correia. A Flow Based Automated Taxi's Route Choice Model Considering Traffic Congestion Delays. (under review)

On-going papers

1. **X. Liang**, K. An, G. H. de A. Correia, and B. van Arem. An optimization model of automated taxis in trip assignment under elastic demand for the first/last mile problem.

2. P. Ashkrof, **X. Liang**, G. H. de A. Correia, and B. van Arem. To foster electric mobility: A review of current initiatives and policies in Europe.

Peer-reviewed conference papers

1. **X. Liang** and G. H. de A. Correia. A Flow Based Automated Taxi's Route Choice Model Considering Traffic Congestion Delays. (under review) Transportation Research Board 99th (TRB 2020), Washington, D.C., US.

2. **X. Liang**, K. An, G. H. de A. Correia, and B. van Arem. An optimization model of automated taxis in trip assignment under elastic demand for the first/last mile problem.

Symposium of the European Association for Research in Transportation (hEART 2018), Athens, GR.

3. **X. Liang**, G. H. de A. Correia, and B. van Arem. Applying a Model for Trip Assignment and Dynamic Routing of Automated Taxis with Congestion: System Performance in the City of Delft, The Netherlands. Transportation Research Board 97th (TRB 2018), Washington, D.C., US.

4. **X. Liang**, G. H. de A. Correia, and B. Van Arem. An optimization model for vehicle routing of automated taxi trips with dynamic travel times. 20th EURO Working Group on Transportation Meeting (EWGT 2017), Budapest, HU

5. **X. Liang**, G. H. de A. Correia, and B. van Arem. Optimizing the Service Area and Trip Selection of Electric Automated Taxis Used for the Last Mile of Train Trips. Triennial Symposium on Transportation Analysis (TRISTAN 2106), Aruba, AW

6. **X. Liang**, G. H. de A. Correia, and B. van Arem. Optimizing the Service Area and Trip Selection of Electric Automated Taxis Used for the Last Mile of Train Trips. Transportation Research Board 95th (TRB 2016), Washington, D.C., US.

7. **X. Liang**, G. H. de A. Correia, and B. van Arem. Optimizing the service zone location and trip selection of electric automated taxis in train trip connections. 18th EURO Working Group on Transportation Meeting (EWGT 2015), Delft, NL.