TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Verkehrstechnik

Taxis, Passengers and Stable Marriage

- Stable simultaneous assignment of taxis to passenger booking requests -

Michal Kümmel, M.Sc.

Vollständiger Abdruck der von der Ingenieurfakultät Bau Geo Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr.-Ing. Gebhard Wulfhorst

Prüfer der Dissertation:

1. Univ.-Prof. Dr.-Ing. Fritz Busch

2. Prof. David Z. W. Wang,
   Nanyang Technological University, Singapore

Die Dissertation wurde am 17.05.2016 bei der Technischen Universität München eingereicht und durch die Ingenieurfakultät Bau Geo Umwelt am 12.12.2016 angenommen.

Acknowledgement

First and foremost, I would like to thank Prof. Fritz Busch and Prof. David Z. W. Wang for their continuous support and willingness to discuss my research at all times. I sincerely appreciate the freedom and support they gave me in choosing and pursuing this topic. Furthermore, I would like to thank all my present and past mentors and colleagues at TUM CREATE, NTU and TUM for their valuable feedback, proofreading and support. I also appreciate the great support from my student assistants. This thesis would not gain its depth and completeness without feedback of about thirty manuscripts reviews and about fifty representatives of taxi and dispatching business ranging from taxi drivers, through dispatching system lead developers to business owners, in Singapore, Germany, several other European countries and USA with whom I had a chance to discuss taxi operation issues and my approach. Lastly, I would like to thank my family and friends for their everlasting encouragement and all the helpful discussions and review rounds, which made this work possible at all.

Kurzfassung

Die vorliegende Arbeit beschäftigt sich mit den Herausforderungen bei der Planung und dem Dispatching von Buchungsanfragen von Taxipassagieren. Den Hauptbeitrag dieser Arbeit stellt die Methodik dar, die Gruppen von Buchungsanfragen von Taxipassagieren Taxigruppen zuordnet und somit die Präferenzen von Taxis sowie Passagieren berücksichtigt. Die entwickelte Methodik wurde für ein zentralisiertes System, welches die Anfragen simultan und in Echtzeit zuordnet, entwickelt. Dies wurde durch die Verwendung eines ‚Stable Marriage' Algorithmus umgesetzt, welcher ursprünglich zur Zuordnung von Männern und Frauen basierend auf deren Präferenzen, entwickelt wurde. Die entwickelte Methodik wird durch ein selbst entwickeltes Simulationsmodell evaluiert. Die Auswertungen ergaben, dass die entwickelte Dispatching-Strategie eine bessere Performance als die üblicherweise verwendeten first-come, first-served Strategie aufweist, wobei die Performance anhand von verschiedenen Indikatoren gemessen wurde (z. B. Wartezeiten der Passagiere und gefahrene Taxikilometer).

Abstract

This thesis deals with the challenge of dispatching taxis to passenger booking requests. The main contribution is the methodology which assigns and re-assigns groups of taxis to groups of passenger booking requests and reflects the interest of not just taxis but also passengers, and which produces stable assignments. The proposed methodology is designed for a centralized system that assigns and re-assigns the bookings simultaneously in decision epochs in real-time. The solution to the stable marriage algorithm, originally designed for matching men and women according to their preferences, is applied to dispatching taxis to passenger booking requests. The proposed methodology is evaluated in a custom-made taxi simulation model. The results indicate that the proposed taxi dispatching strategies outperform a commonly used first-come, first-served dispatching strategy in a number of performance indicators (e.g. passenger waiting time and taxi distance) as well as for various combinations of advance and immediate booking requests.

Table of contents

1 Introduction

Taxi operation is a complex phenomenon characterised by three core aspects: space, time and the amount of supply and demand. Having the right amount of taxi vehicles at the right locations and at the right times is the ultimate objective of taxi operation. This trinity is a paramount of success or failure of taxi operations. It is a constant struggle to balance these three aspects.

Are there too many vehicles if they are not needed? The fleet operating costs rise. Are there too few vehicles? The passengers wait longer or cannot get a taxi at all. Are there just enough vehicles but they are not coordinated? The costs rise and passengers wait. Figure 1 illustrates an example of the over- and under-supply of taxis at one location in one day.

Figure 1 Over- and under-supply of taxis at one location during one day

How can we decrease the mismatch between the taxi supply and demand? Which segments of taxi operation can we influence? How can we improve the way taxis are assigned to passengers? How can we quantify the performance? Are the traditional operational and dispatching rules able to keep up with growing scale of taxi operation? Should taxi dispatchers use other dispatching approaches? There is no universal answer to all these questions.

This thesis attempts to structure problem and, in particular, focus on the dispatching of immediate as well as advance booking requests in order to improve taxi operation by considering not just fleet-related performance but also customer satisfaction and stability of the assignments. The general objective of this thesis is to contribute to further pushing the boundaries of taxi dispatching and scheduling research in order to provide a better service for the passengers and to further improve taxi fleet performance.

Background and development of taxi service

In the early history of taxi service, taxis waited for passengers at demand hotspots, which were typically other transportation hubs such as railway or bus stations, harbours and airports or taxi stands near activity centres. In addition, especially in denser areas, it has become meaningful for taxis to cruise the streets in order to search for passengers. Initially, taxi drivers searched for passengers based on their experience without much coordination among themselves.

Since the passengers cannot actively request a taxi out of the line of sight, some companies introduced taxi booking through phone. In order to disseminate the booking requests to drivers, dispatching centres were established and booking requests were broadcasted by radio. To access the dispatch calls, drivers typically paid a fee (Anderson, 2016). Later, GPS technology enabled the dispatching overview taxi positions exactly. Typically the nearest taxi is dispatched to the incoming booking request. However, especially in demand peaks, the system cannot guarantee there will be an available taxi when a passenger needs it.

Therefore, in order to address this, some taxi operators introduced an option to book a taxi in advance for a specified time (also called pre-booking or advance booking). Advance booking requests, however, introduce additional constraints into the system and are not particularly favourable for system efficiency, as demonstrated in the next chapters.

In a nutshell, the historical developments formed four major market segments, categorised according to how taxi supply is matched with demand: cruising taxis that accept street hails, taxis that wait at stands, taxis dispatched in response to booking requests and taxis with contracts (Frankena and Pautler, 1984) (Salanova et al., 2011) (Aarhaug and Skollerud, 2014). Cruising taxis that accept street hails and passengers from taxi stands serve immediate booking requests only. Taxis under a dispatching system and a contract may serve both immediate as well as advance bookings as shown in Table 1.

Pick-up modes	**Taxi requests**	
	Immediate	**Advance**
Cruising taxis accepting street hails	✓	✘
Taxis waiting at taxi stands	✓	✘
Taxis assigned to booking requests	✓	✓
Taxis with contracts	✓	✓

Table 1 Taxi operation according to pick-up mode and nature of taxi request

Taxi operation can be also categorized according to time eras: pre-dispatching, dispatching and automated era as shown in Table 2. In the pre-dispatching era the taxi drivers are the sole decision makers. Taxi drivers decide where, when and how to search for passengers (how to pre-allocate their vehicles in the expectation of the future demand) and are also responsible for other non-revenue generating tasks such as refuelling, cleaning and servicing. The dispatching era introduces dispatching systems that allow passengers to book a taxi. The dispatching system decides for the taxi driver or in cooperation which taxi will be assigned to pick up which passenger booking request. However, the drivers still have to decide when and where to pre-allocate the taxi vehicles and when to make a break for refuelling, service or maintenance. In the automated taxi era the scheduling responsibility of the driver needs to be replaced by a system.

Tasks	**Pre-dispatching era**	**Dispatching era**	**Automated era**
Assignment of bookings	✘	System	System
Pre-allocation of vehicles	Drivers	Drivers	System
Non-revenue tasks	Drivers	Drivers	System

Table 2 Taxi operation according time eras and control over operational tasks

Despite considerable research efforts, taxi scheduling and dispatching remains a challenge for both taxi operators and dispatchers. Still, focusing on taxis operations is desirable and meaningful; because improving taxi performance may have a higher impact than improving individual vehicle efficiency since taxis generally have a higher mileage than individually owned passenger cars. For example in Singapore in 2014 taxis constitute only about 3% total vehicle fleet size but produce over 14% of traffic on the roads (Land Transport Authority, 2015).

Problem statement

It is difficult to match the taxi supply with demand and, in particular, it is difficult to assign taxis to passenger booking requests. Figure 2 illustrates the increasing complexity of assigning taxis to passenger booking requests. The left column shows an initial situation with the same number of yet unassigned new booking request and taxis, while the right column shows the assignment.

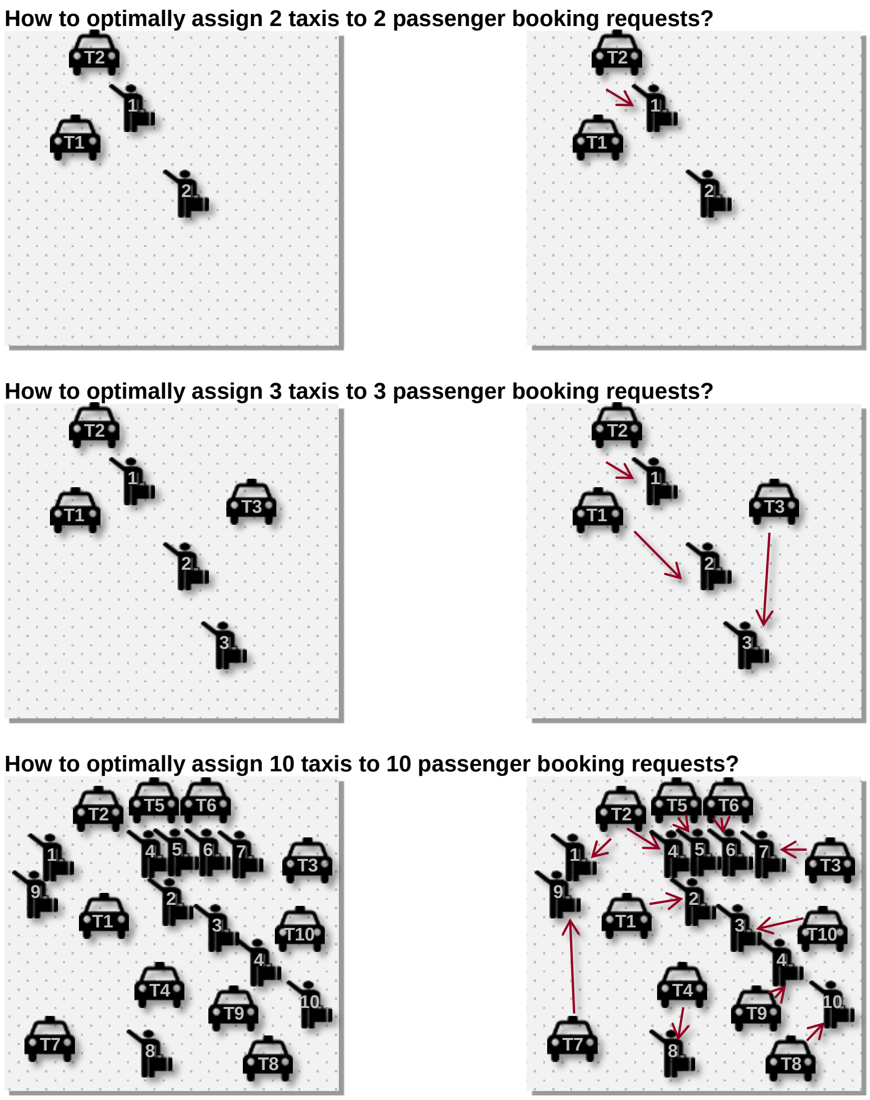


Figure 2 Example of increasing complexity of assigning passenger booking requests to taxis

Humans are able to solve simple instances of taxi assignment to passengers by intuition. Simply by looking at the situation they are able to draw the lines leading to the best performance. Unfortunately with increasing scale this task will soon become overwhelming. The complexity of a simple assignment problem, as illustrated in Figure 2, is O(n!). This makes exact 'brute-force' solutions, which calculate all possible solutions, limited to small scale instances. Note that this example shows only the space dimension of the problem, however, this thesis deals also with the temporal dimension of the problem.

Aims and objectives

This thesis focuses on the problem of assigning taxis to serve booking requests that are coming in real-time. The conventional taxi dispatching approach assigns new booking requests sequentially as they come. This is reliable, but with an increasing scale of taxi operation, increasing number of booking requests and increasing ratio of advance bookings it may become a barrier to a better taxi performance, as illustrated in Figure 13 in chapter 3.1.2. "Fundamental improvement principle".

Therefore, this thesis explores the potential of the fundamental improvement principle in taxi operation – the simultaneous assignment and re-assignment of group of taxis to group(s) of booking requests. Moreover, the assignments reflect the interests of not just taxis but also passengers, which is rarely the case in most approaches. Finally, with human decision makers (taxi drivers) in the loop, it is crucial to produce dispatching assignments, which the individual drivers will follow. Hence, the dispatching system trades optimality for stability of the central control in order to prevent the dispatching system from getting out of control by sequences of individual drivers not following their assignments.

The objectives of this thesis were formulated into the following three research questions:

1. How can the simultaneous assignment of booking requests improve taxi operation?
2. How to prevent schedule from quality degradation over time, especially with high share of advance booking requests?
3. How does accepting individually sub-optimal re-assignments of confirmed bookings affects the system performance?

Also the corresponding hypotheses were formulated:

1. The potential of simultaneous assignment to improve taxi dispatching performance increases with the amount of booking requests that can be assigned at the same time and works the best for immediate bookings.
2. The more advance booking requests, the better the taxi scheduling performance.
3. Greediness of taxis does not pay off. Taxis should respect the system interest.

Contribution

This thesis contributes to taxi research by proposing and evaluating a methodology of dispatching taxis to immediate as well as advance booking requests by reflecting the interest of not just taxis but also passengers and by stable simultaneous group constructive assignments and re-assignments. The proposed methodology is verified in a simulation model in a series of experiments. The simulation is developed and built in MATLAB for the purpose of this thesis, since there is neither a commercial nor available simulation model. This thesis presents proof of the concept that demonstrates that the proposed methodology is able to improve the performance of dispatching taxis to passenger booking requests over the traditionally used approaches.

Furthermore, this thesis thoroughly explores the area of somewhat scattered taxi research literature in order to systematize it. It reveals research needs in taxi research and sketches possible solution approaches. On the side, this thesis also systemizes the method of estimating the costs and profits of individual booking request and handling the schedules and also aims to bring some order to performance indicators. Furthermore, it proposes a framework for the autonomous dispatching of autonomous vehicles. The fundamental limitations of the methodology and validation tools are discussed thoroughly. In addition, the used assumptions about taxi dispatching systems are crosschecked by a survey and interviews with a dozen dispatching system providers (see Appendix 2).

All in all, this thesis proposes a dispatching approach for taxis that demonstrates that especially in large-scale taxi operations there are unused opportunities to serve more customers, to shorten their waiting time and to generate higher profit for the operators and dispatchers at the same time.

Thesis structure

This thesis is structured into a classical blocks as shown by Figure 3. The introduction sets the scene, the research literature is reviewed, three methodological approaches are proposed, the simulation model is described, results of three experiments are presented and discussed, followed by general conclusions and future work directions.

	1. Introduction	
	2. Review of taxi research literature	
Taxi dispatching	Searching, pre-allocation, VRP	Synthesis and lessons learned
	3. Methodology	
Research approaches	Implementation details	Simulation model
	4. Simulation experiments and results	
Experiment 1	Experiment 2	Experiment 3
	5. Discussion and lessons learned	
	6. Conclusion and future work	

Figure 3 Structure of the thesis

2 Review of taxi research literature

Taxi research has focused on three major streams, which all aim to improve the mismatch between taxi supply and demand. (1) In the first stream, the research effort aims to improve the assignment of taxis to passengers (taxi dispatching). (2) The second stream of research uses historical data to predict the future passenger demand and help drivers find passengers and advise them where and when to pre-allocate their vehicles (taxi passenger-searching for on street pickup and pre-allocation strategies). (3) The third stream investigates whether to regulate the taxi market and if so, by which means. The next chapters survey the main research streams in taxi dispatching, passenger searching and regulation, review similar vehicle routing problem and examine the experimental design and approaches of other research studies, summarizes the lessons learned, describe general taxi operation framework, discusses possible solutions and summarize research needs.

2.1 The taxi dispatching problem

In general, real-time taxi dispatching belongs to the class of operational research combinatorial optimization problems. The taxi dispatching problem can be considered as an extended form of a Vehicle Routing Problem (VRP). The classical VRP by Dantzig and Ramser asks "What is the optimal set of routes for a fleet of vehicles to traverse in order to deliver to a given set of customers?" (Dantzig and Ramser, 1959) The classical problem assumes that the demands to transport a commodity, fleet size and vehicle capacities are known before the routing is carried on. Further, it requires the vehicles to begin and end in a central depot, to fulfil all the customer demands, to abide by vehicle capacity constraints and predetermined maximal route length. The overall objective is to minimize the total costs.

The taxi dispatching problem asks: “Which taxi should be dispatched to which passenger booking request?” Taxi dispatching is in fact a scheduling problem and the dispatching is the final act of sending the taxi vehicles to pick up particular booking requests. Scheduling is a process of assigning (and re-assigning) booking requests to taxi schedules according to a scheduling strategy. The scheduling strategy defines which request is assigned to which taxi. Routing is considered to be part of the scheduling process that provides travel times between locations. (Kümmel et al., 2016c)

Terminology and definitions

The following paragraphs introduce and define the terminology used in this thesis, covering the taxi demand, taxi supply and representation of the taxi assignments. The minor terminology is summarized in chapter “Glossary and abbreviations”.

Booking requests / taxi demand

A passenger booking request is a request from passenger to pick up and transport him/her to the desired destination in the desired time. The booking request is understood broadly as the whole process from making a booking enquiry to delivering the passenger to the destination. Booking requests are bounded: (1) in space by their origin and destination locations and (2) in time by the booking time and the desired pick-up time interval.

As an example, Figure 4 depicts the details of passenger booking request 1. The passenger desires to be picked up within the pick-up time interval at the origin and delivered as fast as possible to the destination. From the origin and destination positions, the trip travel time can be estimated. A taxi usually needs some time to reach the origin position of passenger booking, which is called ‘on-call’ time. This time is estimated from the current taxi position and the origin of the trip. Figure 4 depicts a situation, in which the taxi has a short ‘on-call’ time and can reach the passenger at the beginning of the desired pickup time.

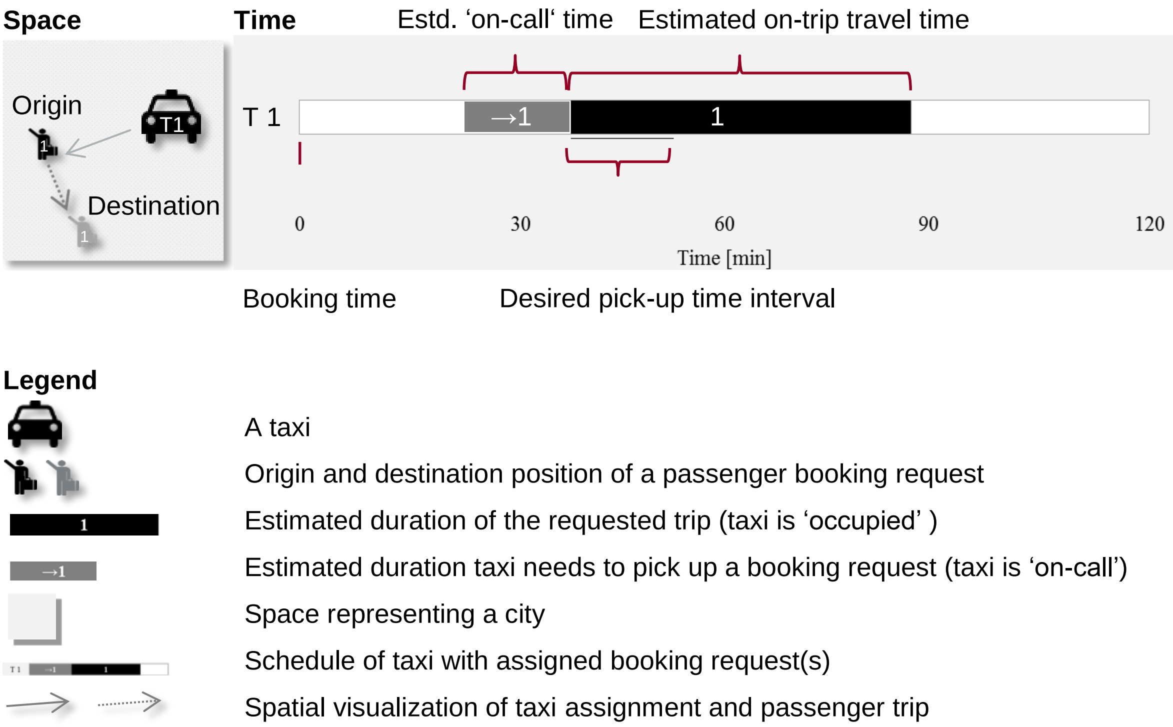


Figure 4 Details of passenger booking request

The passenger booking request may be either immediate or advance depending on the difference between the booking time and the beginning of the desired pick-up time. If there is no difference, the passenger demands the taxi now; the request is called immediate (or current). If there is a time difference, the booking request is called advance (or in advance taxi booking).

The ratio of immediate booking requests to all booking requests is called the degree of dynamism (DOD) (Berbeglia et al., 2010). Weakly, moderately and strongly dynamic systems are distinguished (Larsen et al., 2007).

Taxi drivers and vehicles / taxi supply

Taxis are characterized by their position in time, space and their status. Status of a taxi can be either ‘available’ (also called ‘vacant’), ‘on-call’ to an assigned passenger trip or ‘occupied’ with a passenger trip. When taxi is ‘occupied’, passenger is ‘on-trip’.

Taxi schedule

A taxi schedule is an outcome of scheduling and it is a sequence of booking requests in time for each taxi vehicle. Figure 5 illustrates a schedule and statuses of five taxis (T 1 – T 5) and twelve passengers (1 – 12). The lower part represents the individual schedules of taxis. Each trip with passenger (1 – 12) is preceded by an ‘on-call’ trip to passenger, in which the taxi drives to the origin of a passenger trip (→1 – →12). Upper part of Figure 5 depicts the cumulative taxi fleet status.

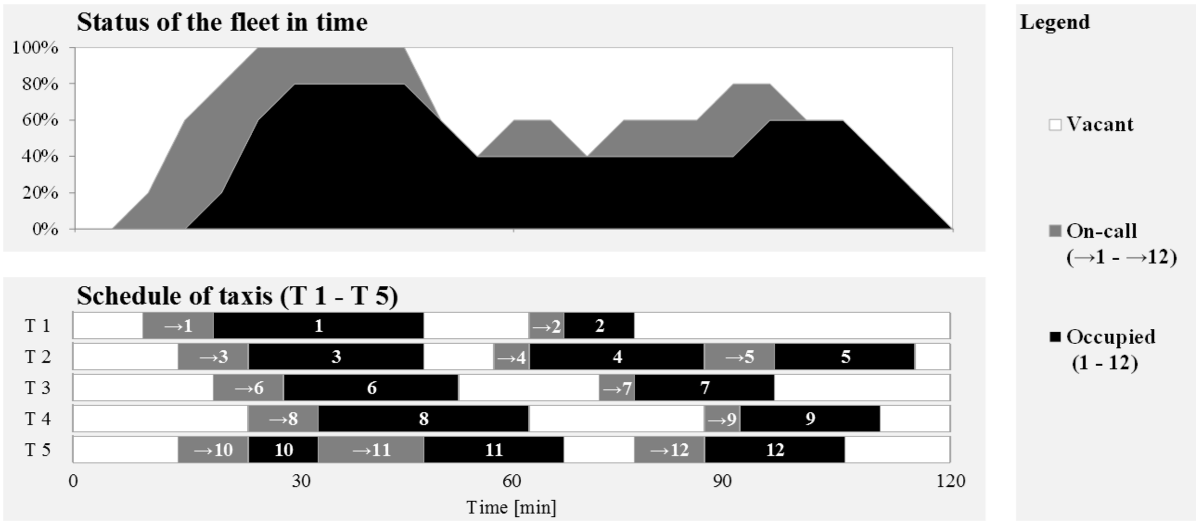


Figure 5 Example of a taxi schedule and status of the fleet in time. (Kümmel et al., 2016a)

2.1.1 Sequential dispatching taxis to immediate bookings

The taxi dispatching problem has been studied from two major viewpoints in the literature: Rule-based sequential (one-by-one) taxi dispatching and simultaneous (concurrent) taxi dispatching. Historically, most of the taxi dispatching research focused on the rule-based sequential assignment of taxis to immediate booking requests. According to this assignment, booking requests are assigned as they come on the first-come, first-served principle to the nearest available taxi as shown in Figure 6.

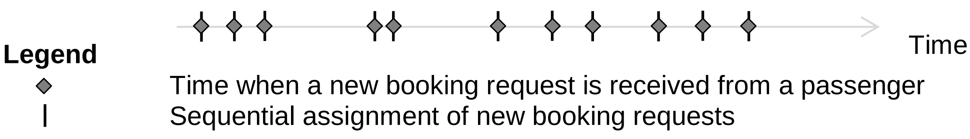


Figure 6 Sequential assignment of new booking requests

One of the first taxi dispatching research papers by Bailey and Clark (1987) investigated efficiency of basic dispatch rules to assign booking requests either to: (1) closest free taxi, (2) closest occupied taxi or (3) the taxi that is ‘vacant’ for the longest time. Later, they investigated whether after delivering a passenger to the destination, the taxi should: (1) return to a centralized base, (2) remain at the delivery location, or (3) relocate to another optimal location based on some relocation algorithm (Bailey and Clark, 1992).

Researchers have investigated details and incremental improvement potential of these rule based dispatching strategies. For example, Lee et al. (2004) and Lee and Wu (2013) concluded that a time-based assignment is superior to Euclidean distance. Indeed, there are many situations, in which the nearest vehicle is on the opposite side of the street, but cannot turn around or cross the street or is delayed by traffic lights.

Most recently, Maciejewski and Bischoff (2015) studied various dispatching rules in under- and over-supply contexts in a realistic, large-scale and detailed simulation in Berlin and Barcelona. They concluded that it is beneficial to change dispatching strategies depending on the supply-demand ratio.

However, none of these incremental improvements of the simple rule based approach were able to overcome a fundamental drawback of the sequential rule based dispatching strategy: Once confirmed, booking requests remain assigned to the same taxi. This may have a detrimental effect on taxi performance over time, as illustrated by the example depicted in Figure 13 in chapter 3.1.2 “Fundamental improvement principle”.

In particular, in high demand situations, the sequential dispatching strategy does not leverage on exchanging some assignments among taxis or assigning booking requests concurrently. As a result, taxi drivers’ costs rise and passengers wait longer or may not get a taxi at all. Table 3 summarizes the objectives and methods of sequential taxi dispatching approaches. (Kümmel et al., 2016c)

2.1.2 Simultaneous dispatching taxis to immediate bookings

In order to overcome the drawback of sequential assignment of taxis to booking requests, some researchers suggested to slightly delay assigning of new booking requests in order to buffer them and assign them simultaneously (concurrently in a group) as seen in Figure 7. Buffering increases the number of assignments possibilities and thus increase the chance of finding more optimal assignment.

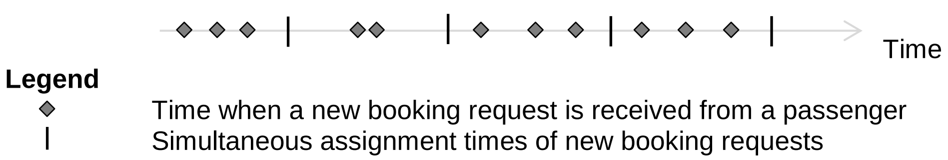


Figure 7 Simultaneous assignment times of new booking requests

Ngo (2004), is one of the first to study simultaneous dispatching taxis to immediate booking requests. He proposed a fuzzy linear framework to model the dispatching rules and combinations (including taxi distance, utilization and costs) and translate them to assignment of taxis to passengers.

Seow et al. (2010) proposed a taxi dispatch architecture to support concurrent assignments of immediate bookings using a decentralized agent group negotiation approach that attempts to minimize the total customer pick-up time of taxis. In every dispatch cycle, taxi agents negotiate on behalf of real taxis how to assign buffered requests. Seow et al. leave the taxi drivers an option to accept or reject the new assignment. An attempt is made to assign all pending requests rejected by drivers in the next dispatch cycle. Moreover, Seow et al. (2010) extended taxi availability to include soon-to-be-available taxis, which allows a larger pool of taxis to be matched.

Zhan et al. (2014) proposed a system for bipartite matching of trips to passengers in New York. They studied a system wide recommendation mechanism that matches taxis and immediate booking requests for the purpose of minimizing taxi costs, the taxi fleet size or combination of both. Passenger bookings are first buffered and then centrally assigned to taxis by solving the minimum weight perfect bipartite matching problem. Their study aims to reduce taxi costs only, and not the passengers' waiting times.

Most recently, Maciejewski et al. (2016) used linear assignment problem solution in order to simultaneously assign new requests to available taxis. They were able to prove real-time running ability on large-scale simulation of Berlin taxi operation. Table 3 provides an overview of the above mentioned approaches to simultaneous taxi dispatching taxis to immediate bookings. (Kümmel et al., 2016c)

2.1.3 Dispatching taxis to advance and immediate bookings

Most of the research studies considered immediate bookings only. However, there are a few research teams, which studied how to dispatch taxis to the combination of immediate and advance booking requests. Most of them employed some form of re-assignment of confirmed booking requests. A review of general dynamic pick-up and delivery problems by Berbeglia et al. (2010) suggests two basic re-assignment strategies: (1) re-assigning some of the confirmed booking requests (schedule adaptation) or creating the schedule anew by re-assigning all booking requests (constructive re-insertion).

Schedule adaptation

The schedule adaptation method is used by Glaschenko and Ivaschenko (2009), who proposed an adaptive scheduling approach with decentralized negotiation of taxi agents in cycles aimed at resolution conflicts of old and new requests. The length of re-scheduling chain is limited by the calculation time needed. The investigated taxi operation handles on average 0.3 requests per taxi per hour and peak demand reaches 0.7 requests per taxi per hour. This approach seems to adapt the schedule at the moment a new request comes. It would be interesting to examine, how this approach would behave under much higher passenger demand and, in particular, what would happen if a request with higher priority will be inserted instead of one with lower priority and no taxi is available for the lower priority one. Moreover, it is to be clarified, how the approach would cope with various degrees of dynamism.

Lee and Wu (2013) also proposed a strategy for dispatching taxis to booking requests. Passenger bookings are first assigned sequentially to the timely-nearest taxi. Later, advance bookings are subject to re-optimization in the 'local planning phase'. The local planning phase either assigns the confirmed advance booking request to another taxi or swaps it with another taxi. Only one booking can be modified per taxi in this phase, or in other words, chains of trips cannot be exchanged. The criteria for successful swap or insertion are (besides feasibility) the costs for the taxi(s). The interests of passengers are not reflected. Their approach uses a tabu set that prevents an advance booking request from being assigned to the same taxi again. The re-optimization is only triggered by the new advance booking request and is done immediately after the initial assignment.

As an exception, Wang et al. (2009) proposed a simple method to dispatch taxis to immediate and advance booking requests that does not re-assign booking requests. In particular, they suggested to chain advance booking requests one after the other and serve these chains by a dedicated part of the taxi fleet and let the other taxis serve immediate booking requests only. Despite partial improvement, this approach does not leverage on the re-assignment of already accepted booking requests. Table 3 summarizes the schedule adaptation approaches in the literature. (Kümmel et al., 2016c)

Constructive re-assignment

In contrast to the schedule adaptation method, which exchanges only some booking requests, the constructive re-assignment (or constructive re-insertion) removes all confirmed booking requests from the schedule and creates a complete new set of assignments. This makes it a static problem that is repeated at regular time intervals. In order to gain time for the static problem computation, these methods typically require ‘freezing’ of new assignments. Therefore, the improvement is calculated either in regular time intervals (also called decision epochs or time slices) or triggered by input data change. (Pillac et al., 2013)

The advantage of this method lies in its ability to cope with many changes at once. The advantage of group assignment is apparent if we consider the situation depicted in Figure 13 in chapter 3.1.2. “Fundamental improvement principle”. By assigning booking requests simultaneously, the total distance to reach all passengers (‘on-call’ distance) can be reduced as compared to sequential assignment. The disadvantage is that the method struggles with lack of continuity and frequent changes. It requires ‘pausing’ of accepting new requests in order to gain time for improving the schedule. To the best of the author’s current knowledge, no taxi dispatching study covers explicitly this approach. (Kümmel et al., 2016c)

	Objective	Method	Authors
Sequential dispatching	Investigate rule based dispatching strategies	Experiments with the model to test the impact of dispatch rules	(Bailey and Clark, 1987)
	Investigate rule based dispatching strategies	A two way analysis of variance with five dispatching strategies	(Bailey and Clark, 1992)
	Combine nearness and utilization in taxi dispatching	Fuzzy rules of 'the nearest vehicle first' and 'the least utilized one first'.	(Shrivastava et al., 1997)
	Investigate impact of real-time traffic	Simulations to investigate the performances of dispatch strategies	(Lee et al., 2004)
	Relax some existing rules for searching 'vacant' taxis	Multi-agent simulations	(Alshamsi et al., 2009)
	Investigate rule based dispatching under various supply	Altering the dispatching rules in the dependence of over or under- supply	(Maciejewski and Bischoff, 2015)
Simultaneous dispatching	Use fuzzy logic to vehicle dispatching	Simulating fuzzy linear framework for linear assignment problem	(Ngo, 2004)
	Let taxi agents negotiate to minimize the total waiting time of a group of customers	Simulating simultaneous (concurrent) assignment multiple taxis in time periods	(Seow et al., 2010)
	Create and evaluate a system wide recommending dispatching system	Simulating bipartite matching of taxis to passengers to increase taxi system efficiency solved by Hungarian method	(Zhan et al., 2014)
	Improve performance of taxi dispatching	Simulating linear assignment problem solution in order to assign simultaneously new requests to available taxis	(Maciejewski et al., 2016)
Schedule adaptation	Re-schedule taxi service before confirming order acceptance in real-time	Schedule adaptation triggered by new event with multi-agent negotiation	(Glaschenko et al., 2009)
	Allow dispatching taxis to advance and immediate bookings	Advance booking trip chaining	(Wang et al., 2009)
	Allow dispatching taxis to advance and immediate bookings	Schedule adaptation with tabu search	(Lee and Wu, 2013)

Table 3 Summary of approaches to taxi dispatching

2.2 Taxi passenger-searching and pre-allocation strategies

Aside from dispatching taxis to the revealed booking requests, taxi drivers pro-actively locate their vehicles where they expect high chances of getting a passenger. Naturally, taxis are attracted to the demand hotpots. Though, taxi drivers may consider the balance between the supply and demand at various times and locations. Table 4 summarizes the discussed searching for passengers and taxi waiting and pre-allocation strategies.

Taxi drivers (at least the perfectly rational deciding ones) estimate and compare possible gains with the costs. The gain can be for example a potential profit (Powell et al., 2011), a function of utility of a taxi stand (Shi et al., 2010) or probability of picking up a passenger en-route and at a taxi stand (Yuan et al., 2011) or imbalance of supply to demand in a region (Miao et al., 2015). The costs can be directly connected with driving there (Shi et al., 2010) and waiting there (Powell et al., 2011).

Mitrović-Minić and Laporte (2004) examined the impact of various waiting strategies effectively postponing the routing assignment decision in a tabu search framework. Kim et al. (2005) researched knowledge building of taxi drivers and implications of information available to taxi drivers and taxi drivers' decisions on the overall quality and operational efficiency of taxi service. They have analysed an effect of a taxi information system that helps drivers efficiently seek passengers represent taxi drivers' destination choice. Shi et al. (2010) also experimented with information provision strategies, which can help drivers to find a taxi stand with waiting passengers. Yuan et al. (2011) proposed a recommendation system, which navigates drivers towards a taxi stand through specific route with respect to current drivers position, however, with no explicit measures to balance supply and demand. Powell et al. (2011) investigated profitability maps overlaid over the geographical ones and their impact to reduce cruising time. In the context of personal rapid transport and balancing the supply of vehicles at each station, Lees-Miller and Wilson (2012) investigated proactive empty vehicle redistribution by sampling and voting. Pavone et al. (2012) investigated rebalancing policies, which periodically assign shared vehicles to passengers waiting at stations with a help of fluid models. However, it does not explicitly account for passenger preferences beyond the wish of getting a taxi, in particular, to get the taxi as fast as possible.

Choo et al. (2012) developed a method for advising taxis to move to regions with a higher demand than supply by issuing advisory tokens to harmonize the pre-allocation. Lee et al. (2013) proposed another recommendation system with no commitments to pick up. They formulated this problem as taxi-customer negotiation in which the ultimate objective of the negotiation is to maximize the global utility for achieving Nash equilibrium in the negotiation process. Bai et al. (2014) noted that in many Chinese cities dispatching systems are not used and taxi drivers have difficulties to find passengers. They suggested addressing this issue by an advisory system, which would be based on principles of stable matching. They compared this system with a virtual vis-a-vis dispatching system, but the results, however, were not promising.

He and Shen (2015) smartphone-based e-hailing applications equilibrium framework by work of Yang et al. (2010). Fagnant and Kockelman (2014) investigated various strategies of pre-allocation strategies of automated fleets. Most recently, Miao et al. (2015) considered pre-allocation based on large-scale real data and the real-time imbalance of supply and demand in dispatch regions. Afian et al. (2015) aimed to estimate unmet demand from the probe data in order to display real-time supply-demand balance.

Most of the pre-allocation, waiting and street pickup methods require some sort of estimation of the future demand. For rebalancing strategies it is typically the expected future passenger demand inflow at the station or taxi stand, for non-restricted cruising applications, it is the expected demand on the link or more commonly an area. While some of the approaches are based on control, the most produce recommendation only.

There are many methods used for the prediction, though one, which outperforms the previous ones, is worth mentioning. Moreira-Matias (2014) proposed real-time prediction method based on machine learning and demonstrated its functionality on the taxi system in the city of Porto. In his PhD thesis he provides detailed review of future demand estimation and pre-allocation approaches. (Moreira-Matias, 2014)

Recharging, service and maintenance are non-revenue generating tasks that are not visible to passengers but are indispensable part of the taxi operation. Recharging or refuelling is an integral part of vehicle operation and especially battery electric vehicles require substantial amount of recharging time on daily basis. Vehicles require regular and on demand inspections, cleaning, servicing, repairs and maintenance. Some of these tasks are very frequent, short and of short term nature (such as vehicle cleaning) and some rather occasional but more time demanding and of long-term nature (such as regular maintenance). The research literature that explicitly takes into account refuelling and service trips is scarce. Fagnant and Kockelman (2014) investigated pre-allocation strategies of automated fleets, considered vehicle refuelling and cleaning at least in a form of simple assumptions of short time interval for refuelling and cleaning after a drop off.

Objective	**Method**	**Authors**
Understand impact of various waiting strategies	Simulate waiting strategies in a tabu search framework	(Mitrović-Minić and Laporte, 2004)
Represent taxi drivers' destination choice	Taxi information system helps drivers efficiently seek passengers	(Kim et al., 2005)
Investigate influence of information on the taxi system performance	Simulation information influence	(Shi et al., 2010)
Reduce taxicab cruising time	Suggesting profitable locations to taxi-cab drivers	(Powell et al., 2011)
Advise taxi drivers where they will find passengers	Recommendation to go towards a taxi stand through specific route	(Yuan et al., 2011)
Investigate rebalancing policies	Fluid model	(Pavone et al., 2012)
Proactive empty vehicle redistribution	Sampling and voting (SV)	(Lees-Miller and Wilson, 2012)
Advisory system to harmonize supply with demand	Dynamically adjust the number of re-allocation tokens	(Choo et al., 2012)
Advise where to find passengers and taxis	A recommender system for finding passengers and 'vacant' taxis	(Yuan et al., 2013)
Negotiate about pre-allocation of taxis to maximize the global utility	Taxi-customer negotiation process with no commitments to pick up	(D. Lee et al., 2013)
Evaluate recommendation system for street hailing as a non-cooperative game between taxi drivers	Assignment based on stable principles to prevent from breaking the assignment	(Bai et al., 2014)
Investigate pre-allocation strategies of automated fleets	Simulate three grid based pre-allocation strategies	(Fagnant and Kockelman, 2014)
Proposing and evaluating several taxi demand predicting algorithms	Evaluating four demand prediction mechanisms	(Moreira-Matias, 2014)
Customers minimize their full trip costs, 'vacant' taxis minimize disutility	Smartphone-based e-hailing applications equilibrium framework	(He and Shen, 2015)
Improving zonal pre-allocation of taxis	Demand-predicting information coupled with real-time sensing information	(Miao et al., 2015)
Display real-time supply-demand balance	Estimate unmet demand from the probe data	(Afian et al., 2015)

Table 4 Review of searching for passengers and taxi waiting and pre-allocation strategies

2.3 Regulation and market issues

Since the 1970s taxi research has focused on answering regulation questions: whether to regulate and if so, how to achieve the regulation objectives. Regulations by controlling market entry were investigated with the help of aggregated and equilibrium models in order to assess their effects and the terminal state of the market (monopolistic, competitive).

Salanova et al. (2011) claims that Douglas (1972) developed the first aggregated model to find out the balance between the regulation and service standards on a cruising (hailing) markets. Many of his followers built on this aggregated framework and extend it by more details.

Cairns and Liston-Heyes (1996) continued in searching for the answers to find out that price regulation is essential and entry regulation useful. Flores-guri et al. (2012) examined the impacts of regulations based on administrative areas. Perhaps the most extensive is the work of Yang and Wong, who pushed the aggregated models to their performance limits by adding a concept of bilateral searching function with frictions, congestion and many other improvements to find out that taxi services should be only subsidized only when there are returns to scale in a Pareto-improving meeting function (T. Yang et al., 2010), (H. Yang et al., 2010), (Yang, 2010), (Sirisoma et al., 2010) and (Yang and Yang, 2011).

Salanova et al. (2011) in their reviews conclude that despite the efforts and useful partial insights into the taxi markets many models were not able to help decision makers to guide their decisions. Table 5 shows the distilled overview of regulation, profitability and market issues discussed above. For more details, an interested reader may refer to the review by Salanova et al. (2011) for profitability, regulation and finding the taxi fleet size issues and aggregated economic models of taxi operations.

Objective	Method	Authors
Find the balance between the regulation and service standards	Aggregated economic model of a cruising (hailing) markets	(Douglas, 1972)
Find out what regulations are essential if any and its effects	Model monopoly, social optimum and the second best situations	(Cairns and Liston-Heyes, 1996)
Characterize bilateral searching and meeting between customers and taxis in order to estimate regulation impacts	Bilateral passenger taxi meeting function with extensions based on equilibrium models	(T. Yang et al., 2010)
Review of the modelling of taxi services with respect to regulation	Survey profitability, regulation and finding the taxi fleet size	(Salanova et al., 2011)
Examine the impacts of regulations based on administrative areas	Aggregated model with zones	(Flores-guri et al., 2012)

Table 5 Review of regulation, profitability and market issues

2.4 Overview of similar vehicle routing problems

The closest related topics to the taxi dispatching are courier, dynamic forms of dial-a-ride problem, stacker crane problem and some warehouse logistics problems. However, the assumptions slightly differ. For example, dial-a-ride assumes sharing of one vehicle by more passengers with different origins and destinations is allowed. Courier routing also allows transporting more items at once. Stacker crane problem, on the contrary, has typically a unit load but not the various degrees of dynamism assumed. Moreover, parcels or containers do not mind waiting so much as human taxi passengers. The author is not aware of a suitable and transferable approach that can be used for dispatching taxis to the combination of advance and immediate bookings.

Shared taxis, dial-a-ride and personal rapid transit

Although ridesharing can provide a wealth of benefits, such as reduced travel costs, congestion, and consequently less pollution, there are a number of challenges that have restricted its widespread adoption. Table 6 summarizes the shared taxis, dial-a-ride and personal rapid transit approaches.

Since 1960s Wilson et al. (1969) investigated a shared dynamic taxi service and routing algorithms, which were evaluated in probably the first simulation model for taxis in FORTRAN. Since then a wealth of vehicle routing problem variations were investigated also sharing vehicle capacity for transporting more passengers or cargo at the same time. The dial-a-ride domains represented by research of Psaraftis (1980), who developed a dynamic programing approach for a single vehicle and immediate requests.

The dynamic version of their dispatching strategy adapts the schedule after new request comes to the system. Horn (2002) proposes a general framework for modelling assignment of on demand vehicles of various purposes including taxis. He used a periodic re-optimization method based on steepest descends involving re-assigning and exchanging trips between itineraries for single as well as shared taxi trips. Santi et al. (2014) quantified potential savings brought by taxi-sharing in a study based on New York taxi data and D'Orey et al. (2012) did similar study for the city of Porto. Atasoy et al. (2015) proposed a concept that blends all sorts of on demand transportation and offers the user a flexibility to choose between travel time, price and comfort. Agatz et al. (2012) reviewed dynamic and distributed taxi-sharing in perhaps the best ride-sharing review to date. Later, Furuhata et al. (2013) surveyed the state of the art of ridesharing and provided useful classifications out the ridesharing aspects.

Objective	Method	Authors
Investigate shared dynamic taxi service	Simulation of different routing heuristic rules	(Wilson et al., 1969)
Further develop taxi-sharing	Dynamic programing approach for a single vehicle and immediate requests	(Psaraftis, 1980)
Propose a framework for modelling assignment of on demand vehicles	Periodic re-optimization method based on steepest descends	(Horn, 2002)
Evaluate dynamic taxi-sharing	Distributed algorithms	(D'Orey et al., 2012)
Review of taxi-sharing	Critical review of known approaches	(Agatz et al., 2012)
Review the state of the art of ride-sharing	Classification of ride-sharing element	(Furuhata et al., 2013)
Evaluate taxi-sharing benefits	Construction of share-ability networks	(Santi et al., 2014)
Offers user the flexibility to choose between travel time, price and comfort	Integrate assortment optimization, vehicle routing, scheduling and pricing	(Atasoy et al., 2015)

Table 6 Review of shared taxis, dial-a-ride and personal rapid transit

Express mail delivery, courier and logistics

Table 7 highlights some express mail delivery, courier and logistics approaches. Benyahia and Potvin (1997) investigated genetic programing for real-time dispatching of express mail delivery. Gendreau et al. (1999) investigated a tabu search heuristic with adaptive memory to improve solution of local express courier service problem. Gendreau et al. (2006) applied neighbourhood search with ejection chains and tabu search. The neighbourhood is based on ejection chains (Glover, 1996), where a request is moved from one route to another and may eject a request from that route. The ejected request is then moved to yet another route. The chain ends when the insertion of a request in a route does not lead to any ejection. A chain might be of any length and might be cyclic or not. A constrained shortest path problem is defined and solved to find the best possible ejection chain in the neighbourhood. Ichoua et al. (2006) extended the tabu search to exploit knowledge about future demands for real-time newspaper delivery.

Chang and Yen (2012) developed a demand-node merging procedure to reduce the size of courier services problems. The research by Ferrucci et al. (2013) in the context of routing of urgent goods studied neighbourhood search with a tabu and a variable neighbourhood structure. Furthermore, Ferrucci and Bock (2014) evaluated tabu search on a real network with disruptive events. Ehmke and Campbell (2014) elaborate the topic of customer acceptance in the urban delivery context and pointed out by that 'a customer perception' creates the service. Treleaven et al. (2012) explored algorithms for pick-up and delivery problems and, in particular, a connection between Euclidean Bipartite Matching Problem and the theory of random permutations. They have presented a polynomial-time, asymptotically optimal algorithm for the stochastic stacker crane problem. Very interestingly Gath et al. (2014) explored more flexible and adaptive algorithms for the same day delivery business. They have developed a nested Monte-Carlo search with a policy adaptation algorithm in conjunction with objective functions. Recommendable review papers are by Berbeglia et al. (2010) and Prodhon and Prins (2014).

Objective	Method	Authors
Investigate real-time dispatching of express mail delivery	Genetic algorithms to minimize operations costs and maximize service quality	(Benyahia and Potvin, 1997)
Improve solution of local express courier service problem	Tabu search heuristic with adaptive memory	(Gendreau et al., 1999)
Improve solution of local express courier service problem	Neighbourhood search with ejection chains and tabu search	(Gendreau et al., 2006)
Exploit knowledge of future demands	Ejection chain extension	(Ichoua et al., 2006)
Review of on-demand vehicle algorithms	Critical review and summary of existing notable approaches	(Berbeglia et al., 2010)
Explore connection between bipartite matching problem	Asymptotically optimal algorithm for stochastic stacker crane problem	(Treleaven et al., 2012)
Reduce the size / CPU time of solving a courier problem	A demand-node merging procedure	(Chang and Yen, 2012)
Improve routing of urgent goods	Tabu search with a variable neighbourhood structure	(Ferrucci et al., 2013)
Account for disruptive events in courier delivery	Tabu search on a real network	(Ferrucci and Bock, 2014)
Review of location-routing problems	Critical review and summary of existing notable approaches	(Prodhon and Prins, 2014)
More flexible and adaptive algorithms	A nested Monte-Carlo search with a policy adaptation algorithm	(Gath et al., 2014)

Table 7 Review of express mail delivery, courier and logistics

Automated taxis and on-demand vehicles

Replacement of conventional taxi vehicles by automated ones will challenge taxi operation in many ways. Notably, taxi drivers' decision making will be substituted by an integrated scheduling system that also manages the non-revenue generating trips such as vehicle pre-allocation, recharging and service trips.

The research related to autonomous vehicle scheduling is hidden under occasionally surprising key-words; for example: Shared autonomous vehicles, also known as aTaxis (Zachariah et al., 2014), personal rapid transit (in an open-control framework) (Berger et al., 2011), flexible mobility on demand systems (Atasoy et al., 2015), demand responsive transport services (Diana, 2006), demand responsive transport systems (Deflorio, 2011), taxi on demand (Thomopoulos et al., 2007), driverless public transport pods, mobile location-based services (Silva and Mateus, 2003), taxicab networks (Zhang and He, 2012), unmanned automated vehicles, mobile robots, cyber cars (Awasthi et al., 2011), smart cyber fleets (Billhardt et al., 2014), cybernetic transportation system (Wang et al., 2008), autonomous dial-a-ride transit (Dial, 1995), tele bus or autonomous free-floating car-sharing fleets (Firnkorn and Müller, 2014).

Other vehicle routing problems

For more reviews of dynamic vehicle routing and dispatching by Gendreau and Potvin (1998), review of partially dynamic vehicle routing by Larsen et al. (2002), taxonomic review of vehicle routing problems by Eksioglu et al. (2009), review of dynamic pick-up and delivery problems by Berbeglia et al. (2010), a review of dynamic vehicle routing problems by Pillac et al. (2011) and (2013) reviewed related to express pick-up and delivery processes by Ferrucci and Bock (2014), recent survey on location-routing problems by Prodhon and Prins (2014) as well as the most recent survey on dynamic and stochastic routing problem by Ritzinger et al. (2015) and review of simulating demand-responsive transportation Ronald et al. (2015).

Other more distantly related problems general include pick-up and delivery problem, dynamic (real-time) (fleet) routing problem, paratransit problem, (dynamic) (centralised) fleet management problem / system, cybernetic transportation, one-way car-sharing systems, mobile server applications, robotic load balancing. (Gendreau and Potvin, 1998) (Xin and Ma, 2004) (Awasthi et al., 2011) (Pavone et al., 2012) (Pavone et al., 2012) (Billhardt et al., 2014) Here is at least a summary of most commonly used solutions methods used for vehicle routing problem: (1) exact approaches: branch and bound, branch and cut, (2) constructive heuristics: savings: Clark and Wright, matching-based, multi-route improvement heuristics, (3) two-phase algorithm heuristics: cluster-first, route-second algorithms, the petal algorithm, the sweep algorithm, route-first, cluster-second algorithms and (4) metaheuristics: ant algorithms, deterministic annealing, genetic algorithms, simulated annealing and tabu search. ("Solution Methods for VRP," 2013)

Experimental design of research studies

The literature review also offers interesting insights into the objectives, performance indicators and size of the experiments. These, at first glance seeming similar, show in fact a wide array of various approaches and ways of tackling a similar problem from various viewpoints, as summarized in the following chapters. Table 8 provides an overview of the objectives of selected taxi studies. It groups the objectives into increase or maximize, into minimize or decrease or reduce and into ensure categories. In the right column the respective authors are listed.

Objectives		Authors
Maximize / Increase	Taxi profitability / profit	(Bailey and Clark, 1987) (Bailey and Clark, 1992) (Glaschenko et al., 2009) (Shi et al., 2010)
	Service quality / service level	(Bailey and Clark, 1987) (Benyahia and Potvin, 1997) (Glaschenko et al., 2009)
	The number of dispatched calls	(Alshamsi et al., 2009)
	Global utility	(D.-H. Lee et al., 2013)
	Total ridership	(Horn, 2002)
	The number of satisfied drivers and riders in the system	(Agatz et al., 2012)
	Efficiency of dispatch system in handling taxi bookings	(Lee et al., 2004)
Minimize / Decrease / Reduce	Fleet costs / taxi costs / operations costs	(Ngo, 2004) (Lee and Wu, 2013) (Zhan et al., 2014) (Maciejewski et al., 2016)
	Customer waiting time	(Benyahia and Potvin, 1997) (Bailey and Clark, 1992) (Lee et al., 2004) (Alshamsi et al., 2009)
	Empty taxi cruise time	(Alshamsi et al., 2009) (Powell et al., 2011)
	Total vehicle travel time / system-wide travel time	(Horn, 2002) (Agatz et al., 2012)
	Total vehicle travel distance / system-wide vehicle-miles	(Horn, 2002) (Agatz et al., 2012)
	Total waiting time of a customer group	(Seow et al., 2010)
	Taxi idle time	(Alshamsi et al., 2009)
	Taxi fleet size	(Zhan et al., 2014)
	Taxi fleet costs and the taxi fleet size at the same time	(Zhan et al., 2014)
	The number of vehicles performing rebalancing trips	(Pavone et al., 2012)
	Weighted combination of the time to service all customers and of the total degree of 'dissatisfaction	(Psaraftis, 1980)
Ensure	Relatively fair distribution of orders	(Glaschenko et al., 2009)
	Every station should reach an equilibrium between excess vehicles and the demand	(Pavone et al., 2012)

Table 8 Overview of the objectives of selected taxi studies

The most commonly used key performance indicators of a wide array of key performance indicators in three categories related to passengers, taxis and dispatching systems are summarized in Tables 9, 10 and 11 respectively. These tables cluster similar indicators together, such as for example that the 'average passenger waiting time' and 'the relative reduction of average passenger waiting time' is considered to represent the same passenger waiting time notion. In addition, various terms are clustered together, so that for example 'passenger', 'customer' and 'client' waiting time are considered synonyms.

Category	Taxis key performance indicators	Authors
Size	The number of taxis required to satisfy all trips / taxis used / total number of vehicles / the taxi fleet size / the number of vehicles required to provide a given level of service / average number of in-transit vehicles	(Wilson et al., 1969) (Kim et al., 2005a) (Wang et al., 2009) (Pavone et al., 2012) (D.-H. Lee et al., 2013) (Bai et al., 2014) (Zhan et al., 2014)
Profit	Fleet costs / taxi costs / operations costs / taxi operation cost / total costs	(Ngo, 2004) (Shi et al., 2010) (Lee and Wu, 2013) (Zhan et al., 2014) (Miao et al., 2015) (Maciejewski et al., 2016)
	Taxicab profit / profitability	(Bailey, Jr and Clark, Jr, 1987) (Bailey and Clark, 1992) (Glaschenko et al., 2009) (Shi et al., 2010) (Powell et al., 2011)
	Taxi revenue	(Shi et al., 2010)
	Empty taxi trip cost	(Zhan et al., 2014)
Time	'Vacant' cruising time / empty taxi cruising time	(Alshamsi et al., 2009) (Shi et al., 2010) (Seow et al., 2010) (Powell et al., 2011)
	'Vacant' time / (total) taxi idle time	(Kim et al., 2005a) (Alshamsi et al., 2009) (Zhan et al., 2014)
	Estimated and actual travel time to pick up / 'on-call' time / average pick-up trip time	(Lee et al., 2004) (Ngo, 2004) (Maciejewski et al., 2016)
	Total time / total vehicle travel time / system-wide travel time	(Horn, 2002) (Powell et al., 2011) (Agatz et al., 2012)
	'Vacant' waiting time	(Shi et al., 2010) (He and Shen, 2015)
	Travel time per trip / occupied cruising time	(Horn, 2002) (Shi et al., 2010)
Distance	Taxis average total mileage / total vehicle travel distance / system-wide vehicle-miles	(Horn, 2002) (Agatz et al., 2012) (Bai et al., 2014)
	'On-call' distance / distance to passenger	(Ngo, 2004) (Bai et al., 2014)
	Total empty trip distance / average total idle distance	(Zhan et al., 2014) (Miao et al., 2015)
	Occupied travel distance	(Kim et al., 2005a)
	Taxi 'vacant' mileage	(Bai et al., 2014)
Other	The occupancy rate / taxi utilization ratio (occupied / service) / taxi time utilization rate / taxi time utilization time	(Shi et al., 2010) (Lee and Wu, 2013) (D.-H. Lee et al., 2013) (He and Shen, 2015)
	Taxi queue length at a taxi stand	(Shi et al., 2010)
	The ratio of unoccupied to total drive time	(Maciejewski et al., 2016)

Table 9 Key performance indicators of selected taxi research studies – passengers

Category	Passengers key performance indicators	Authors
Waiting time	(Average) passenger waiting time / (average) customer (client) waiting time / relative reduction thereof	(Bailey and Clark,1992) (Benyahia and Potvin, 1997) (Lee et al., 2004) (Kim et al., 2005a) (Alshamsi et al., 2009) (Seow et al., 2010) (Shi et al., 2010) (Choo et al., 2012) (D.-H. Lee et al., 2013) (Lee and Wu, 2013) (Bai et al., 2014) (He and Shen, 2015) (Maciejewski et al., 2016)
	Customer queue length at a taxi stand / expected passenger queue	(Kim et al., 2005a) (Shi et al., 2010)
	Customer waiting time at individual taxi stand	(Shi et al., 2010)
	95th percentile of client waiting time	(Choo et al., 2012)
	Clients with over 60 min wait	(Choo et al., 2012)
	Total number of waiting customers	(Pavone et al., 2012)

Table 10 Key performance indicators of selected taxi research studies – taxis

Category	Dispatching system key performance indicators	Authors
Service	Service quality / service level	(Bailey, Jr and Clark, Jr, 1987) (Benyahia and Potvin, 1997) (Kim et al., 2005a) (Glaschenko et al., 2009)
No. of served requests	No of served trips / the number of dispatched calls / total ridership / the number of completed bookings / average total number of calls served by each cab / percentage of served calls	(Horn, 2002) (Kim et al., 2005a) (Alshamsi et al., 2009) (Powell et al., 2011) (Lee and Wu, 2013) (Zhan et al., 2014)
	The number of unsuccessful bookings / percentage of cancelled calls (cancellation rate) / the number of lost orders	(Alshamsi et al., 2009) (Glaschenko et al., 2009) (Lee and Wu, 2013)
Response time	Average dispatch time	(Alshamsi et al., 2009)
	Delayed pick-ups	(Glaschenko et al., 2009)
	Urgent order average response time	(Glaschenko et al., 2009)
Taxi-sharing	Shared trips	(Santi et al., 2014)
	Saved travel time	(Santi et al., 2014)
	Saved trips	(Santi et al., 2014)
	Mean value of the ratio of service time to direct driving time taken together with the extreme worst service time	(Wilson et al., 1969)
	Average number of vehicles performing rebalancing	(Pavone et al., 2012)
Pre-allocation	Supply demand ratio of whole city and region and mismatch thereof	(Miao et al., 2015)
	Computational solution times	(Psaraftis, 1980) (Horn, 2002)
Other	Orders allocated automatically	(Glaschenko et al., 2009)
	Fleet utilization effectiveness	(Glaschenko et al., 2009)
	Relatively fair distribution of orders	(Glaschenko et al., 2009)
	Expected travel time on a link	(Kim et al., 2005a)

Table 11 Key performance indicators of selected taxi research studies – dispatching system

Similarly, the sizes of simulation experiments used in the research studies are listed in Table 12. The experiments are sorted according to the fleet size, but they also feature the number of passenger requests, ratio of requests per taxi and per hour, duration of the experiment and network data where available.

Fleet size	Requests	Reqs./taxi/h	Duration	Data	Authors
1	<= 10				(Psaraftis, 1980)
8	6000				(Shrivastava et al., 1997)
4	160				(van Hentenryck et al., 2010)
12	140	1.9	6 h		(Benyahia and Potvin, 1997)
20	0-3000				(Ngo, 2004)
20-90					(Wilson et al., 1969)
10-100					(Atasoy et al., 2015)
22				Singapore	(Wang et al., 2009)
25-40					(Bailey and Clark, 1992)
50	2000				(Pavone et al., 2012)
60	5000				(Atasoy et al., 2015)
100	1000				(Bai et al., 2014)
100-250	520	2-5	2 h		(D. Lee et al., 2013)
220	4282	1.6	24 h		(Horn, 2002)
280-450					(He and Shen, 2015)
500	800-8000	0.64-6.4	2.5 h	Singapore	(Lee and Wu, 2013)
500				San Francisco	(Miao et al., 2015)
538					(Powell et al., 2011)
800		0.41 – 1.65	3 h		(Shi et al., 2010)
1000		1-4			(Seow et al., 2010)
1622	43348	5.35			(Alshamsi et al., 2009)
2000				Singapore	(Lee et al., 2004)
>2000	>13000	0.3-0.7			(Glaschenko et al., 2009)
5700	27376 – 136880	0.3-1	24 h	Berlin	(Maciejewski et al., 2016)
12000			70+40 d		(Yuan et al., 2011)
13586					(Santi et al., 2014)
15000			1 w	Singapore	(Choo et al., 2012)
32368 – 33999	489234 – 524792	0.6	3*20 min	New York	(Zhan et al., 2014)

Table 12 Overview of the size of experiments of selected taxi research studies

2.5 Synthesis and the lessons learned from the review

This chapter synthesises and summarizes the lessons learned from the literature review. The general dependencies in a taxi ecosystem are sketched and research needs in the taxi dispatching research literature are summarized.

2.5.1 General framework of taxi operation

Taxi operation can be described as a black box with general inputs and outputs. Figure 8 visualized the general inputs such as taxi supply, taxi demand, environment and strategies that match the supply with demand.

Taxi supply is in general, constituted by taxi vehicles and taxi drivers. The intersection of available taxis and available taxi drivers defines the available taxi fleet. Various strategies, such as shift planning, pre-allocation of the vehicle and general scheduling, as well as random effects such as unexpected delays define the taxi spatial-temporal availability.

Taxi demand consists of the number of potential customers and their individual desires to be transported. Perhaps the most important individual considerations that define the actual demand are: From and to where the individual likes to be transported, when ideally, when the latest, on which costs and whether it necessary to book a taxi, or just wait for one passing by.

The environment, in which taxi supply can match with taxi demand, is constrained by a network of roads with topologies, travel times and road tolls. Other variables such as price of labour, vehicles, energy to fuel them and the fares also shape the taxi supply and demand matching. In general, taxi operation and business strategies worldwide have many flavours and there are many models employed, from self-employed drivers with own vehicles to pure employees acting on behalf of a dispatcher.

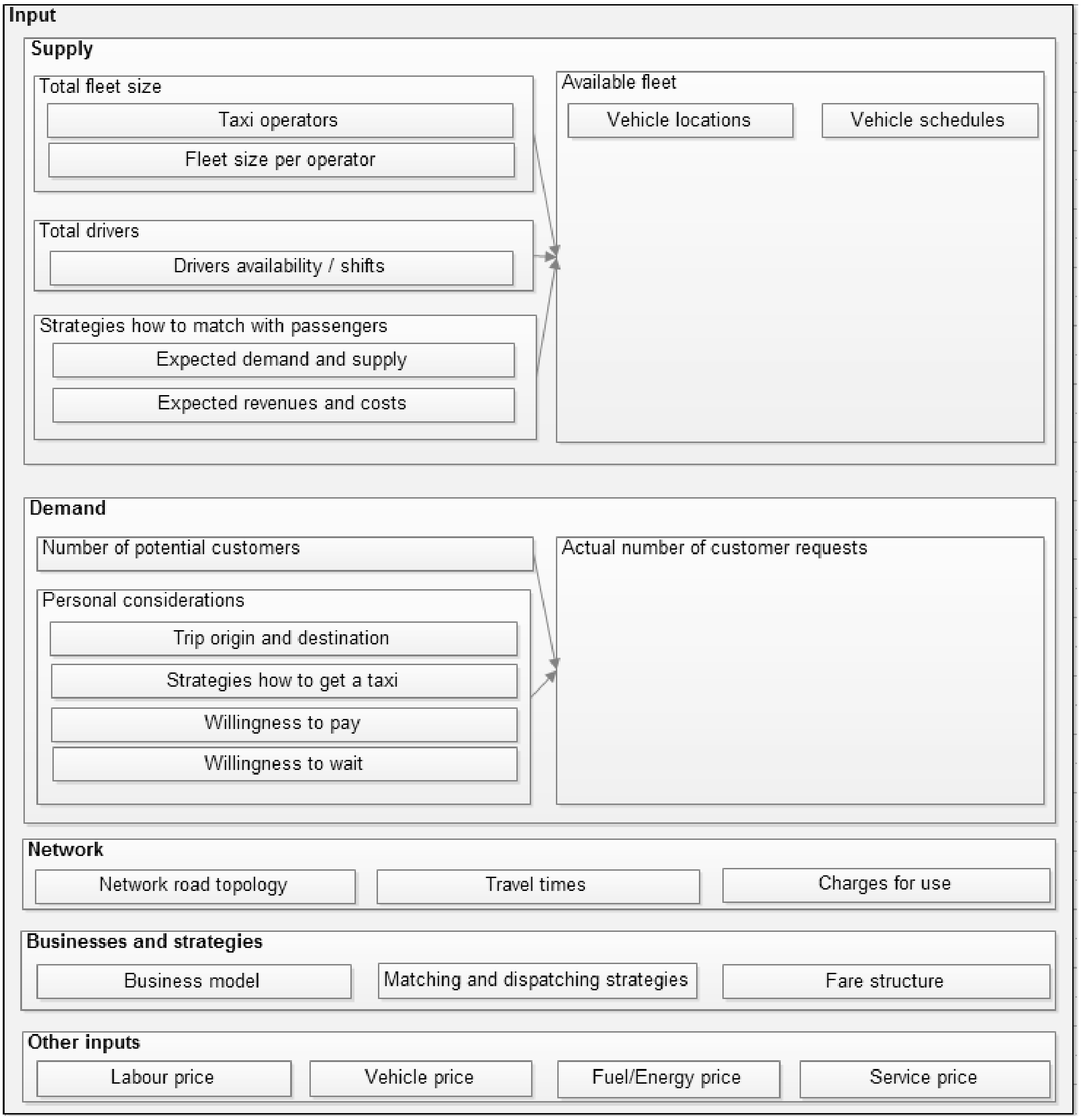


Figure 8 General dependencies in taxi ecosystem – input

Figure 9 visualize general outputs of taxi operation is the service, its quality and corresponding financial reward. Essentially, the number of passengers transported (satisfied demand) and the waiting time and costs (service quality) are relevant to passengers, while the revenues and costs from the corresponding taxi actions are relevant for taxi drivers and operators.

Essentially, there are three types of revenue and costs: Based on distance, time and events. The issue of revenue and costs are discussed in detail in the section 3.5 Performance indicators analysis.

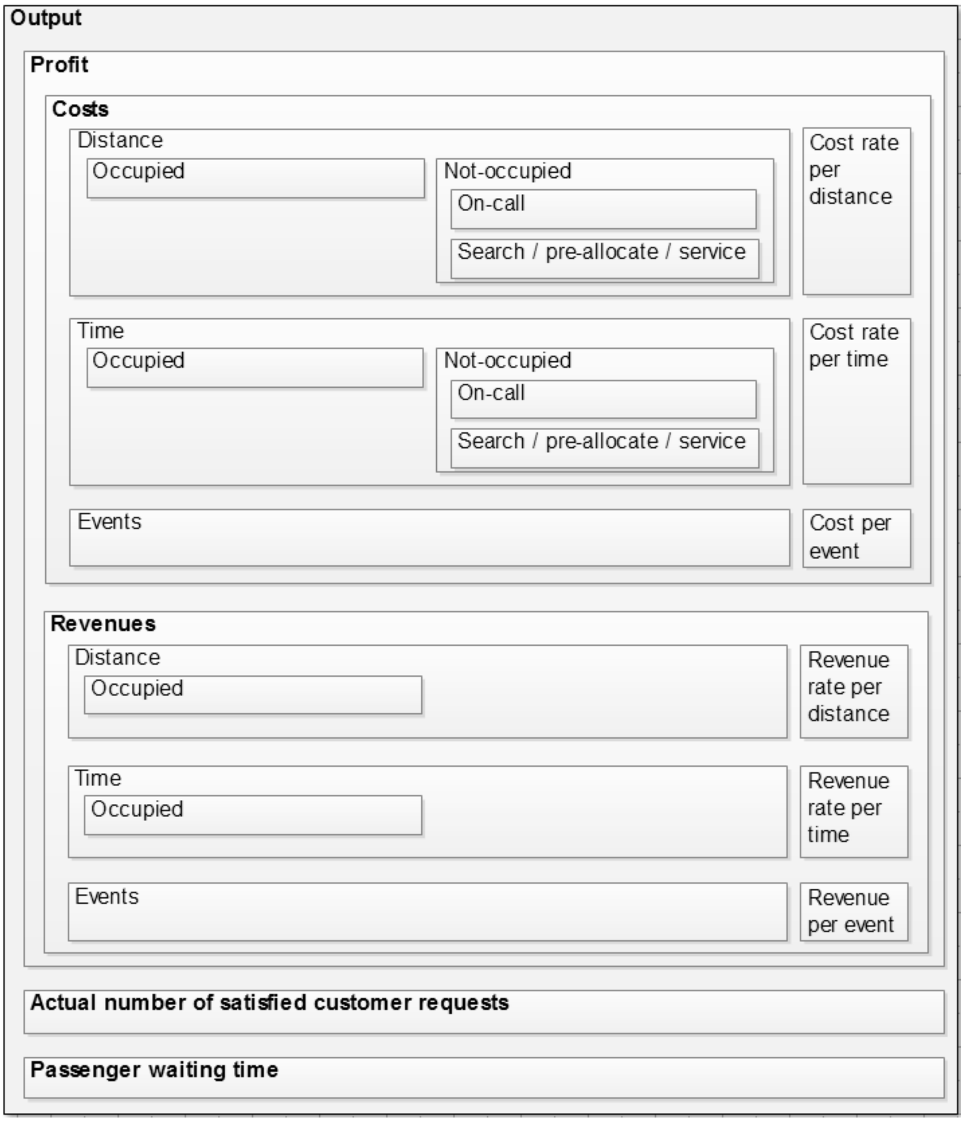


Figure 9 General dependencies in taxi ecosystem – output

2.5.2 Analysis of possible solution approaches

As mentioned earlier, the fundamental problem of taxi operation is the mismatch between the supply and demand. This mismatch may be manifested in dimensions of space, time, the amount of the taxi supply and demand, fares and other aspects. The mismatch impacts the costs and profit of taxis as well as passengers waiting time and the ability to get a taxi.

In order to improve the mismatch between taxi demand and supply, for instance one can increase the supply. This can be done by simply buying more taxis or by issuing more taxi permits (medallions) on regulated markets. As a result, the average ratio of supply to demand improves (on condition the demand remains the same), together with fixed costs, but that alone actually not necessarily solve the mismatch problem. The essential condition for improving the mismatch is to fulfil the trinity of having the right amount of taxis at the right positons and at the right time. Simply, adding more taxis is not sufficient.

Alternatively, the demand for taxis could be reduced. Passengers can be for example deferred by increased fares or just by providing a really terrible service. Again this leads to increasing the ratio of taxi supply to taxi demand, but not necessarily solving the mismatch problem.

One can propose to increase the taxi availability to ask the drivers to drive at least a certain distance each day, in order to circulate the taxis. This solution may improve the situation for on-street pick-ups, but it increases costs, can be difficult to enforce and chiefly, similarly as before it misses the target of ultimately solving the mismatch problem.

It seems that it also makes sense to try to improve the matching by pre-allocating taxis to locations with favourable supply-demand ratio. Individual drivers do this naturally based on individual experience and knowledge. Pre-allocation may increase profit from serving otherwise unserved passengers. On the other hand, it requires creating predictions about future supply-demand balance and coordination strategies and increases costs. Wrong predictions or unexpected competitor moves may result in even higher costs. Though, in general, this approach seems quite promising to prepare the ground for the actual matching and, to certain extent, it improves the mismatch.

Other solutions like sharing taxi rides are proposed to decrease the mismatch. Technically sharing actually reduces the demand. But, similarly to the solution with reducing the demand, this solution per se does not hit the point. Moreover, taxi-sharing has, except potentially cheaper rides and potentially less vehicles used, many disadvantages. In particular, the passenger total travel time may increase, and the passenger may have to wait for potentially delayed fellow passengers.

The next option is to improve the matching by improved dispatching taxis to booking requests. Even with the simple algorithms, the dispatching systems, in general, may increase taxi availability, reduce unnecessary costs and reduce passenger waiting time.

Another solution is to extend the search space by increasing the pool of taxis that participate on dispatching. Arguably, the bigger is the vehicle fleet, the higher is the chance that there will be a taxi nearby where it is needed and may reduce waiting times and costs. Table 13 lists the above mentioned solutions together with their main pros and cons. This thesis aims to look at the most promising single solution from these, which is improving the dispatching strategies.

Possible solutions to taxi supply demand mismatch	Pros	Cons
Increase supply Add more taxis, issue more taxi permits	+ Increases supply to demand ratio and thereby chances of finding available taxi	- Increases fixed costs - Might not necessarily solve the mismatch problem
Decrease demand Increase fares, decrease service quality	+ Increases supply to demand ratio and thereby chances of finding available taxi	- Increased price for example may defer some customers but does not necessarily improve the matching per se
Increase supply availability Ask drivers to drive at least a certain distance daily or work longer	+ May increase taxi availability mainly for on-street pick-ups	- Increases costs - May not necessarily help to solve the mismatch - May be difficult to enforce and control
Pre-allocate taxis Improve the matching by pre-allocating taxis to locations with favourable supply-demand ratio	+ To certain extent individual drivers do this naturally based on individual experience and knowledge + Aggregating the knowledge and coordinating pre-allocation of taxis + Pre-allocation may increase profit from serving otherwise unserved passengers	- Requires creating predictions about future supply-demand balance and coordination strategies - Pre-allocation increases costs - These predictions may turn wrong, causing extra costs
Share taxis Improve the matching by sharing trips	+ Many studies demonstrate beneficial effects of using less vehicles to transport more passengers + Passenger costs per ride decreases	- The passenger total travel time may increase - Passenger may have to wait for potentially delayed fellow passengers
Improve dispatching strategies Improve the matching by improved dispatching taxis to booking requests	+ May increase taxi availability + May reduce unnecessary costs + May reduce passenger waiting time	- Requires strict cooperation
Increase matching search space Improve the matching by increasing the pool of taxis that participate on matching	+ One taxi is subject to several dispatching systems. + May reduce waiting times and costs	- Simultaneous assignment of new bookings or re-assignment of confirmed requests and the resulting benefits may be prohibitively difficult in reality

Table 13 Considered solutions to improve taxi supply demand mismatch

2.5.3 Research needs in assignment of taxis to passengers

This section summarizes the main research needs in assignment of taxis to passengers. Further research and evaluations of new approaches addressing these concepts and gaps would be fruitful.

Simultaneous (re-)assignment of booking requests is not commonplace

Algorithms for assigning taxis to passenger requests in reality might not be utilizing the full potential of technology and scale of taxi operation, since most approaches do not consider simultaneous assignment and re-assignment of booking requests.

Dispatching taxis to any combination of advance and immediate bookings

Dealing with the combination of immediate and advance booking requests seems to be a significant challenge for taxi service providers. Similarly, typically once accepted advance booking requests assignments are fixed and not re-assigned to be served by other vehicles.

Fundamental drawback of schedule adaptation approach

The experiment of Lee and Wu (2013), which can be considered as the most elaborated taxi dispatching approach in taxi research literature to date, fully revealed a drawback of the schedule adaptation approach on large-scale taxi systems. Especially in high demand systems and in peak times, where there is the highest potential for re-assignment of bookings, the frequency of incoming booking requests is the highest. That means, there is the least time for computing re-assignments when it is needed the most and the problem complexity is the highest because of the highest number of possible assignment combinations. In particular, Lee and Wu (2013) assumed that the booking rate arrival in the experiment equals 800-8000 bookings per hour, out of which 50% are booked in advance. This creates an average time gap for re-assignment that is only 0.5 – 4.5 s long. Assuming that in Singapore, in which they located their experiment, there is about one million trips, out of which 10% are booked in advance, the average time gap for re-assignment would be 0.9 s. This is very short time given that the presence of sensing, positioning and communication delays in the real world.

Furthermore, the schedule adaptation method may work well if the degree of dynamism is close to 0 (most of the bookings are known in advance and only some new bookings trigger re-assignment) but it is unclear how it will perform under degrees of dynamism close to 1 (almost all bookings are immediate). (Kümmel et al., 2016c)

Interests of both taxis and passengers

The assignment of the majority of the above listed approaches focuses only on the interests of fleet operators. Ngo (2004), Wang et al. (2009), Seow and Lee (2010), Lee and Wu (2013), Zhan et al. (2014) Pavone et al. (2012) do not explicitly account for passenger preferences beyond the wish of getting a taxi, in particular, to get the taxi as fast as possible. Passengers do mind waiting for a taxi. However, only a few studies considered not just the interests of taxis but also the passengers. Alshamsi et al. (2009) for instance utilize the multi-agent systems advantage and assume that the passengers time should be minimal, taxi ‘vacant’ distance and idle time minimal and the dispatchers number of dispatched calls maximal.

Stability of assignments and re-assignments

The stability of produced assignments are also not explicitly considered in the dispatching strategy. This is typically overcome by providing recommendations only, which the drivers may not follow or reject if not suitable for the driver from any reason such as in case of Seow and Lee (2010).

Stability consideration is important for the systems with human taxi drivers. Understandably, drivers may not from the local perspective understand, why they are asked to pick up a passenger which is further away, which would have been better from the system perspective. Hence, under the stability consideration, the driver-passenger pair does not have an incentive to switch and therefore more likely to keep the assignments. This may seem trivial, however, the consequences of not following system optimal assignments may be as follows. If assignments are frequently broken, they must be re-optimized, so that other taxi(s) will have to pick up the confirmed requests. This has several implications: The dispatching system must be able to re-assign already confirmed bookings. As the literature review suggests, only some approaches are re-assigning them.

In an extreme case, the not following the assignments may lead to spiralling effects and precluding the dispatching system to effectively control the assignment of new bookings to taxis. In summary, the robustness of the control algorithm (dispatching system) must be considered.

However, the stable solutions are ‘paying a price’ for the stability. Stable solutions are known to perform worse than the optimal ones. Though, it seems still worthwhile to investigate these, in particular, for their robustness, which prevents chains of uncontrolled reactions initiated by non-compliance of some taxi drivers with the assignments.

Growing scale of taxi operations

Growing scale of taxi operations means also growing challenges and opportunities. In particular, the naïve sequential based algorithms may cause some unnecessary pick-up costs, as the scale of taxi dispatching grows. For larger problems, there should be other algorithms.

Dispatching / pre-allocation not connected with other parts of taxi operation

Mostly, the pre-allocation depends on taxi drivers. Very seldom; there is a central entity that advises drivers where to pre-allocate their vehicles. And if so, usually it is not the part of the dispatching system package. Though, it would be worthwhile to actually connect these two operation issues into one.

Automated vehicles and their automated operation and management

To the best of author's knowledge, the existing approaches in taxi literature do not explicitly consider all tasks of the automated vehicles and are focusing almost exclusively on scheduling of transportation requests. Furthermore, the existing approaches typically propose either centralized or decentralized scheduling but rarely combine the advantages of both. Lastly, the existing approaches seldom consider integration of the infrastructure and service providers in a transparent manner. A separate satellite study at the end of the methodology chapter tackles this issue and prepares a ground for the automated scheduling of automated on demand vehicles.

Lack of algorithms that can learn from own mistakes or changed conditions

Kocer and Arslan (2012) as one of the few investigated ability of solution algorithms to learn from past experience. In particular, they see the gap of existing approaches, because they do not use past solutions to solve current problem quicker and to find better solution with less computation. The exploited 'transfer learning' approach and proposed a routing strategy based on genetic programing combined with learning. Automation of the whole scheduling process may also be an opportunity for implementation of machine learning algorithms, instead of traditional heuristics. Machine learning algorithms may have tremendous advantage in the amount of data being able to process and the ability to learn from mistakes of all the vehicles in the past.

Balancing key performance indicators

Taxi dispatching is a connected ecosystem and it is possible to optimize the taxi dispatching towards various goals and the interests of various parties. Optimizing towards only selected parameters may lead to seemingly good results, which, however, may worsen other unreported parameters. Specifically, chapter 3.5 "Performance indicators analysis" proposes a small fix to the current state, in particular, a method how to logically structure the performance indicators and relate them to each other.

Research knowledge transfer to industry

Lesson learned from the practice operation and from the literature: The research progress, in particular in matching, assignment and routing algorithms, does not seem to be reflected in the taxi dispatching reality. As pointed out by Cordeau et al. (2002) most algorithms lack some attributes essential for application in the real operation in terms of accuracy, speed, simplicity and flexibility. The complexity of taxi scheduling is difficult to seize, taxi simulation models limited by simplifying assumptions and the gap between research of new algorithms and its application is wide. Moreover, to the best of knowledge of the author there are no truly easy-to-use simulation tools that allow testing new dispatching strategies before implementing them.

Dealing with advance passenger booking requests

Common sense suggests two approaches to manage advance taxi bookings: (1) to assign advance bookings sequentially (the same way as immediate bookings) and to confirm advance bookings but buffer them and (2) to assign them shortly before the requested pick-up time as if they were immediate bookings.

If advance booking are assigned as they come, the respective taxi has the responsibility to serve the booking in the future. However, this has several shortcomings and limitations for the particular taxi. For instance, the taxi driver should not accept any other bookings that would prevent him/her from serving already confirmed booking(s). In order to do so, the driver must estimate the likelihood of being able to pick up the potential and confirmed passengers. This involves very precarious and difficult decision-making, which must be done quickly and safely. Consequently, they may either have long breaks before a pick-up (which cost time and money) or not pick up the confirmed passenger at all (which may be penalised by the dispatching system).

Assignment of advance booking shortly before the requested pick-up time has also drawbacks. Especially in demand peaks, there may not be enough taxis to serve the confirmed advance booking requests. Then taxis from further away must be dispatched, which on one hand increases waiting time of passenger and on the other hand, fleet operation costs increase. In order to reduce and ideally avoid such situations, a human dispatcher(s) must intervene and warn or coordinate the respective taxis. As a consequence, this may further increase fleet costs.

Difficulty of drawing conclusions from other simulation studies

This is a small example of an effort to summarize the impact of some of the most relevant research papers on taxi dispatching in terms of profit or costs and waiting time changes as compared to their respective benchmarks. In addition, partly due to the combination of the gaps mentioned before, it is very difficult to reasonably compare similar studies, since either the results are lacking some indicators, they use different benchmarks or, in general, are incomparable due to various different supply, demand, environmental and dispatching system assumptions. Figure 10 provides evidence thereof.

Proposed algorithm	**Profit / costs change**	**Waiting time improvement**
(Ngo, 2004)	Difficult to compare with FCFS	Difficult to compare with FCFS
(Zhan et al., 2014)	Cost reduction up to 94%	Not provided
(Seow et al., 2010)	Not provided	Up to 42% for 20 vehicles
(Maciejewski et al., 2016)	Not provided	Almost the same until demand level 3
(Glaschenko et al., 2009)	Not provided	Not provided
(Wang et al., 2009)	Not provided	Not provided
(Lee and Wu, 2013)	Not provided	Difficult to compare with FCFS

Figure 10 Illustration of results of the most relevant taxi dispatching research approaches

Lack of taxi research benchmarks

The FCFS dispatching strategy is not the most elaborate one, however, it is the most common one in the taxi industry. Moreover, it is commonly used in benchmarking of taxi research papers because unlike others it is relatively easy to replicate. A commonly accepted benchmark or benchmarks in taxi operation research (as it is common in general vehicle routing problems) together with lists of best solutions is desirable, but to the best of the knowledge of the author this is not yet the case in taxi research. Therefore, for the time being, it makes sense to stick with using the simple FCFS strategy for the sake of comparison and replicability.

3 Methodology

In order to answer the three research questions a corresponding methodology is developed in the course of the thesis. This section summarizes major assumptions and the fundamental improvement principle, proposes three approaches to taxi dispatching and scheduling, elaborates details thereof, presents a path towards fully autonomous vehicle dispatching of autonomous vehicles and describes the simulation model and key performance indicators used.

3.1 Research approaches

The three approaches are used to answer the following research questions:

1. How can the simultaneous assignment of booking requests improve taxi operation?
2. How to prevent schedule from quality degradation over time, especially with high share of advance booking requests?
3. How does accepting individually sub-optimal re-assignments of confirmed bookings affects the system performance?

Table 14 summarizes the key properties of the three proposed approaches. In particular, it highlights the assumed degrees of dynamism, the proposed method for assigning new bookings and re-assigning already confirmed ones and other key features. Figure 11 then presents a graphical summary of all three proposed approaches to taxi dispatching. These are the same figures as presented in the three approaches but displayed together, in order to highlight the differences among the three approaches.

All three approaches begin with an existing taxi schedule. Approach 1 is based on simultaneous insertion of the new requests. For Approach 2 and 3 the insertion is done sequentially, because it is faster, and the requests will be re-assigned anyway in the re-assignment stage. Though, it is important to first assign new requests, to quickly respond to the customers, confirming that there will be a taxi for them. Approach 1 ends after the assignment phase. Approaches 2 and 3 follow with the re-assignment phase, in which all or some of the already confirmed bookings are re-assigned. If the schedule after re-assignment is better than the existing one, it is accepted. The whole process repeats in the next decision interval.

	Approach 1 Simultaneous assignment of taxis to new immediate bookings	**Approach 2 Re-assignment of all confirmed immediate and advance bookings**	**Approach 3 Re-assignment of the timely nearest confirmed immediate bookings**
Assignment of new bookings	Simultaneous	Sequential	Sequential
Re-assignment of confirmed bookings	No	Simultaneous	Simultaneous
Degree of dynamism	1	<0,1>	1
Schedule concerned	All new requests are simultaneously assigned	All bookings are re-assigned	Only the timely nearest bookings are re-assigned
Greediness of players in re-assignment	Not considered	Not considered	Greedy trips and optionally greedy taxis

Table 14 Overview of the three approaches

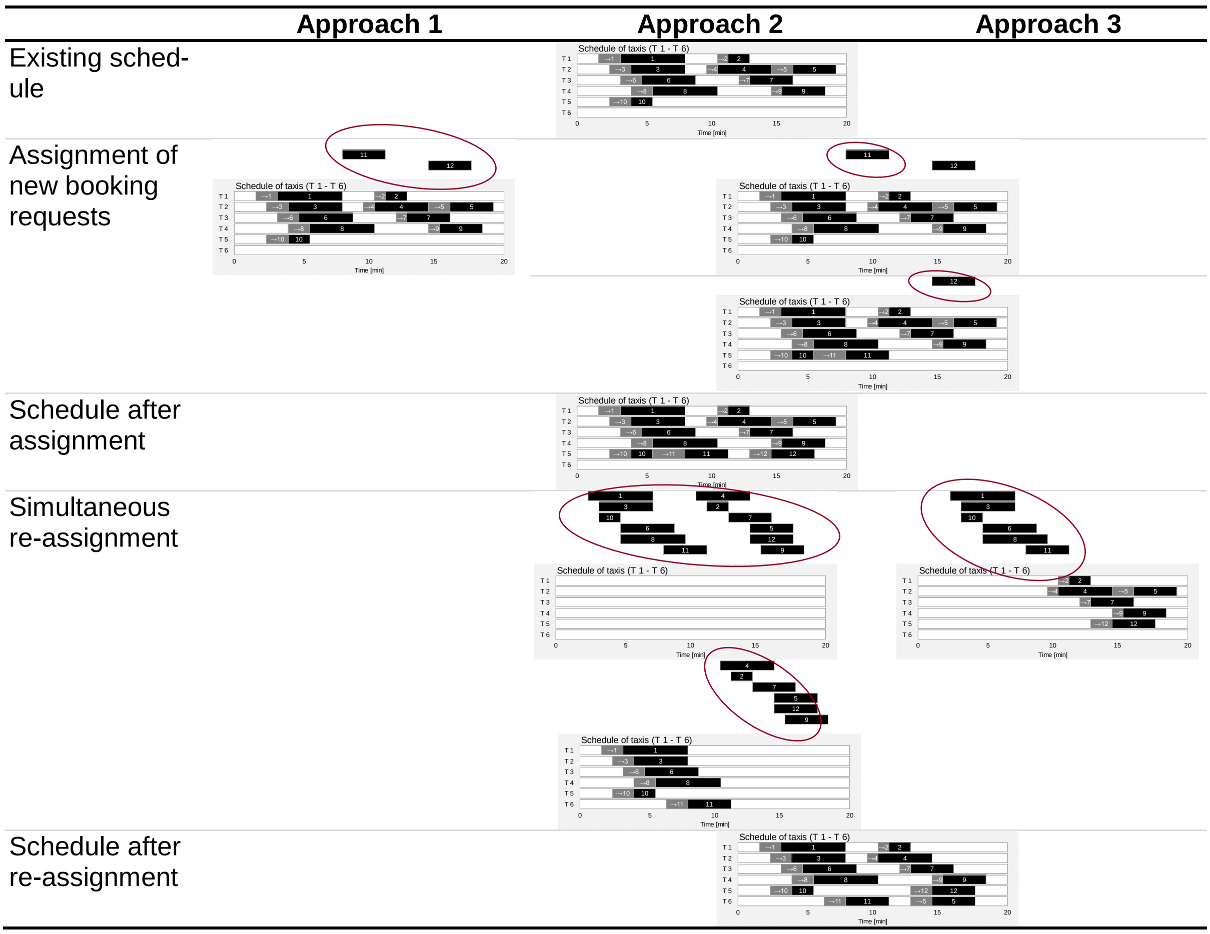


Figure 11 Graphical summary of all three approaches

3.1.1 Major assumptions

The following considerations are used in formalizing the problem. The dispatching strategy is proposed for coordinated taxis with a central dispatching of taxis in response to booking requests. Passenger booking requests are revealed to the dispatching centre. The dispatching centre either confirms or rejects the request depending on vehicle availability. Once confirmed, the passenger request must be served.

The passenger booking requests are defined by their origin and destination positions, time of booking and the requested pick-up time interval. Passenger requests cannot be postponed to be served later, unlike for example in the study by Angelelli et al. (2009). Moreover, passengers are not willing to share the taxi with anyone else unlike in the dial-a-ride problem and are willing to wait for taxis only limited amount of time.

The taxi fleet size remains unchanged during the simulation. Taxis drive only when responding to a call or with a passenger on board. Taxis do not break down. Taxis wait at their initial locations or drop-off location of the last served passenger for their next assignment. Taxis do not use any form of anticipation of future requests. Taxis are geographically distributed at locations where the shifts begin. Thus, the dispatching problem can be called: MDMVCDVRPPDTWDCR – Multi Depot Multiple Vehicle Capacitated Dynamic Vehicle Routing Problem with Pick-up and Delivery Time Windows and Deniable Customer Requests.

The passenger is required during the booking process to input the desired origin, pick-up time and the destination. This is necessary especially for advance booking requests in order to make sure that the newly assigned trip will not cause a conflict with another already scheduled passenger trip. A passenger, who books a taxi, initially receives just the booking confirmation that there will be a taxi. The passenger receives the particular taxi ID shortly before the pick-up time. This is done in order to allow constructive re-assignment until the latest moment in order to improve the schedule. The assumed taxi booking process is sketched in Figure 12.

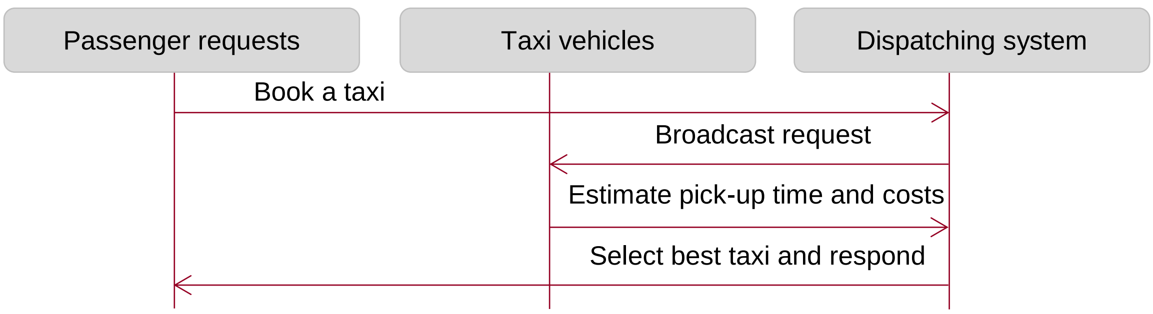
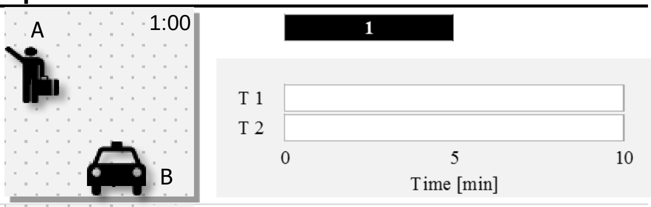
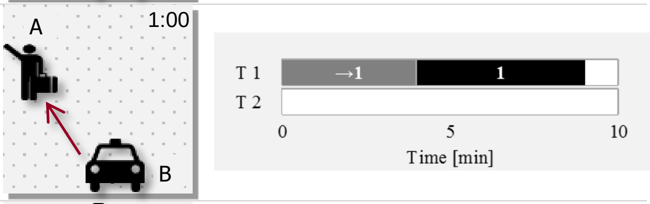
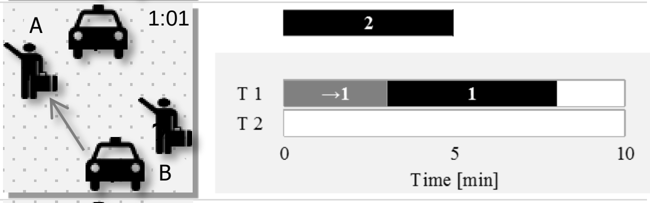
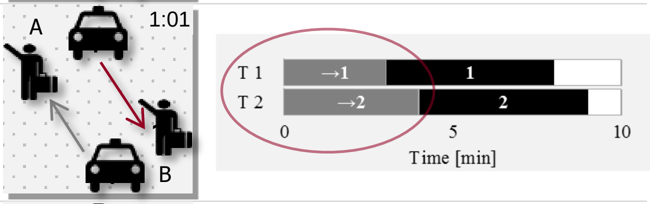
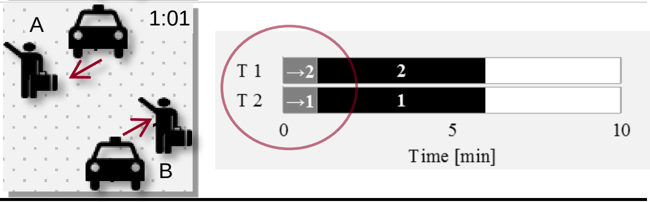


Figure 12 Assumed taxi booking process description

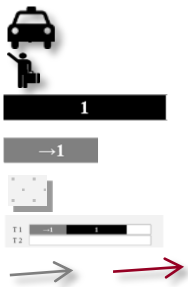
3.1.2 Fundamental improvement principle

Traditionally, new bookings are assigned on a first-come, first-served principle (FCFS), a very intuitive approach to taxi dispatching. Once new booking is confirmed by the dispatching system, it remains in the schedule of the particular taxi. That, however, may lead to a degrading of scheduling quality over time, as illustrated by the example in Figure 13. In particular, the FCFS strategy does not utilize the synergic effects of exchanging some assignments. As a result, taxi drivers drive further than necessary and work longer hours for less income, and passengers wait longer or may not get picked up at all.

In the example in Figure 13 (a) 'Passenger 1' at position A requests a taxi. Figure 13 depicts the events in both space and time. The desired pick-up location, the desired pickup time and the expected duration of the trip are depicted in Figure 13. The nearest available 'Taxi T1' at position B is assigned to pick up 'Passenger 1'. In Figure 13 b this is shown by the arrow, which is pointing now from 'Taxi T1' to 'Passenger 1' and by assigning the trip with 'Passenger 1' to the schedule of 'Taxi T1'. Taxi 'on-call' time is labelled as →1. Then 'Passenger 2' requests a taxi to transport him/her from position B (Figure 13 c). Now the nearest available 'Taxi T2' is at position A. Consequently, both taxis have long drives, in the opposite direction along the same route, to pick up their passengers (Figure 13 d). They would benefit from exchanging their assignments as highlighted in Figure 13 e by the red bars. Notice that after the swapping of assignments the length of 'on-call' time reduces for both taxis and both passengers get the taxi sooner. (Kümmel et al., 2016c)

Description	Space / Time
(a) Passenger 1 makes an immediate taxi booking at 1:00 PM. He desires to be picked up at position A as soon as possible and reach his destination as fast as possible. The trip duration is estimated to 5 min and displayed above the schedule.	
(b) Dispatching system finds the nearest available taxi, confirms the booking to passenger 1 and dispatches taxi T1. Taxi T1 is now 'on-call' to passenger 1, as displayed by an arrow. Estimated time of arrival of taxi T1 to passenger 1 is 4 min from now.	
(c) At 1:01 PM Passenger 2 makes a booking request. He desires to be picked up at position B as soon as possible The trip duration is also 5 min. Meanwhile, taxi T2, which just dropped of a passenger, becomes available.	
(d) Dispatching system finds the nearest available taxi (T2), confirms the booking requests to passenger 2 and dispatches taxi T2. Taxi T2 is now 'on-call' to passenger 2 as displayed by an arrow. Estimated time of arrival is 4 min from now.	
(e) Consequently, both taxis have long drives, in the opposite direction, to pick up their passengers. They would benefit from exchanging their assignments. Taxis can reduce 'on-call' times and passengers waiting times.	

Legend



Available taxi
Origin position of booking request
Estimated pick-up time and trip duration of a booking request
Estimated 'on-call' time of a taxi to pick up a booking request
Space representing a map of a city
Schedule of taxis with assigned booking request(s)
Spatial visualization of taxi assignments

Figure 13 Illustration of fundamental improvement principle (Kümmel et al., 2016c)

3.1.3 Simultaneous assignment of taxis to immediate bookings

In general, the goal of the first approach is to propose a dispatching strategy, which assigns taxis to immediate passenger booking requests simultaneously. In order to address the research needs identified in chapter 2.5.3 “Research needs in assignment of taxis to passengers”, the assignment should also be done in a stable manner to guarantee robustness of the dispatching system and should consider the interests of taxis and passengers at the same time.

The status quo first-come, first-served approach assigns booking requests sequentially. Especially under high passenger demand, with many bookings near each other in space and time, this strategy cannot leverage on the opportunities to assign some of the bookings more suitably. This leads to unnecessary taxi moves to pick up passengers and indirectly to reducing the service capacity and quality.

Therefore, the first approach studies the possibility of simultaneous assignment of passenger bookings. That means that this dispatching strategy assigns a group of taxis to pick up a group of passengers. The principle advantage is that unnecessarily long pick-up trips are reduced and therefore the total costs reduced and passenger satisfaction increased. The principle disadvantage is that the requests must be buffered for some time in order to gather a group together and to create a ‘freeze’ period, in which the group assignment is calculated.

Simultaneous assignment is not enough of a new contribution since there are a handful of simultaneous assignment approaches in the literature. However, to the best of authors knowledge, they do not consider the preferences of passengers and stability of assignments. Why are these important and how do they shape the Approach 1?

First of all, the objective of many studies is to reduce the costs of the taxi fleet (expressed in various forms). In order to do that, taxis are allowed to negotiate amongst themselves using fuzzy logic or linear assignment, without considering the needs and the interests of the passenger. This may have an impact on the revenue.

Secondly, stability is a not common concept for taxi dispatching. A central dispatching entity, which tells the drivers which passenger to pick up, should produce assignments that the drivers are likely follow. Imagine, what may happen in a system, which does not support drivers’ compliance. If the driver do not follow the given assignment and picks another passenger instead (e.g. during a street pick-up), the dispatching system still must assign another taxi or reorganize assignment of a number of taxis in order to pick up the confirmed passenger. For example, in San Francisco in 2007 more than half of the assignments were not followed and the passenger with confirmed booking did not get a taxi. (City and County of San Francisco Taxi Commission Public Convenience & Necessity Report, 2007). Not following the assignments, may lead to customer frustrations and a cascade of more drivers disobeying the assignment and potentially spiral-like trajectory of the dispatching system. For more robust systems, one should consider systems, in which the players do not have an incentive to change the given assignments.

In summary, the governing principle of the proposed methodology to improve taxi dispatching is the simultaneous assignment and stable assignments. The proposed dispatching strategy leverages on the fundamental principle of simultaneous assignment of many booking requests to taxis. This is done in regular time intervals in real-time. The simultaneous assignment is a principle, which allows improving the performance, by matching many booking trips at once. Figure 14 demonstrates the basic principle using an example of an existing schedule of six taxis, to which two new booking requests with ID 11 and 12 are added simultaneously.

Simultaneous assignment of new booking requests

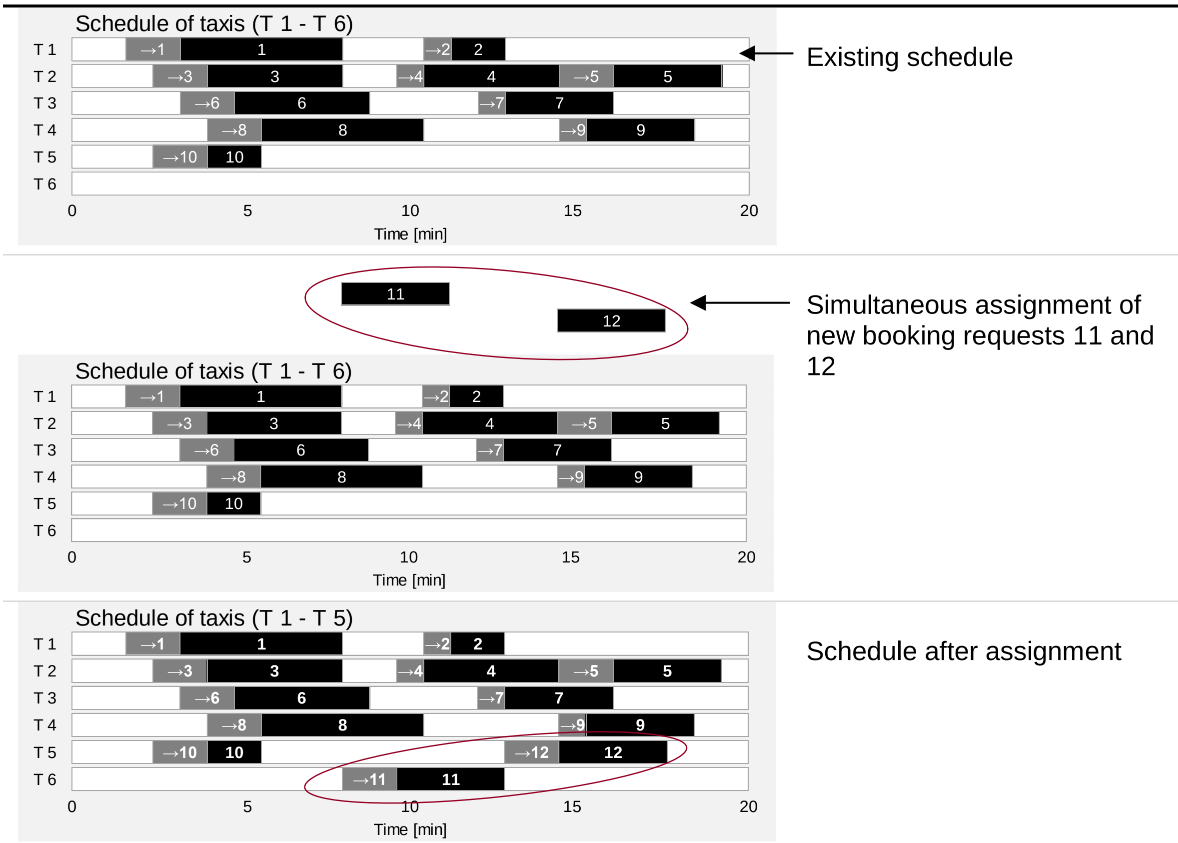


Figure 14 Example of simultaneous assignment of new booking requests

One of the few studies that explicitly consider the interest of passengers, despite not focusing on taxi dispatching but street hailing segment, is the study by Bai et al. (2014). They noted that in many Chinese cities, dispatching systems are not used and taxi drivers have difficulties to find passengers. They suggested addressing this issue by an advisory system, which would be based on principles of stable matching. They compared this system with a virtual vis-a-vis dispatching system. The results, however, were not promising. Despite an average not-occupied taxi mileage improvement, the average total mileage and the average passenger waiting time increased. Nonetheless, their approach inspired this study to apply the stable marriage algorithm in the taxi dispatching context and answer some of the research needs in order to estimate the potential of this algorithm for the taxi dispatching problem. (Kümmel et al., 2016d)

The following paragraphs uncover three major factors in implementation of the proposed dispatching strategy, which were not yet addressed and propose three solutions to handle them. First, Bai et al. (2014) assume in their street hailing study that passengers are willing to wait for an infinite amount of time until they are served by a taxi. This study aims to overcome this unrealistic simplification by introducing a constraint, which limits passengers' willingness to wait. If the limit is exhausted, the passenger will seek an alternative transport option.

The introduction of the willingness to wait, however, has an implication for the original stable marriage algorithm, which assumes that everyone has a preference about every one of the opposite party. If the taxi is not able to pick up a passenger within the pick-up time window, the passenger in fact cannot have a preference. This leads to the stable marriage problem with incomplete preference lists that is described by Brito and Meseguer (2006). The preferences of taxis and passengers are constituted in the following fashion such as by for example Bai et al. (2014). Passengers prefer a taxi that can pick them up first; taxi drivers, on the other hand, prefer booking requests with the least costs.

Next, generating assignments in regular time intervals from buffered requests might cause incoherence in the assigned pairs of taxis and passengers from one interval to another. For illustration, imagine a taxi on its way to the passenger with confirmed booking when a new passenger request appears closer to the current positon of the taxi and the taxi is reassigned to pick up the closer one. Undoubtedly, taxi drivers would prefer to break the commitment and to pick up the closer one, but with no taxi left to pick up the formerly confirmed passenger, passengers would be frustrated. The remedy is to ensure that the confirmed trips are really picked up and to forbid taxis with 'On-call' status to participate on matching process.

The last concern, mentioned but not addressed by Bai et al. (2014), relates to the length of the decision epoch. The ultimate objective is to balance two conflicting interests: On one hand to maximize the number of buffered requests for simultaneous assignment and on the other hand to confirm the bookings to the passenger as fast as possible. As the length of the buffering interval increases, the number of passenger and taxis the matching pool also grows, which increases the potential for improvement.

There are theoretical and practical limits to set the decision epoch length. Firstly, if the length of the decision epoch is higher than the willingness to wait, the passenger cannot get an answer the booking request. Secondly, from the practical user perspective, any buffering strategy holds the requests before distributing them back to the customers. This implies that the passenger must wait for the first reply (presumably longer time compared to FCFS strategy). On the other hand, the system response time is known (defined by the length of the decision epoch) and the total waiting time from the moment of booking until the arrival of the taxi can be shorter. However, this benefit is not seen in the first moment and people may be eager to receive an instant reply. Therefore, with respect to this rationale, this study deduces that the appropriate length of the decision epoch should be really fast. (Kümmel et al., 2016d)

3.1.4 Re-assignment of all confirmed immediate and advance bookings

The objective of the approach is to find a scheduling methodology, which will improve the traditionally used first-come, first-served strategy for any combination of advance and immediate booking requests. The approach should overcome the fundamental drawback of the schedule adaptation approach and as well serve the interests of both taxis and passengers and assign them in the stable manner for robustness.

The fundamental challenge is the re-assignment of already confirmed booking requests so that the dispatching does not fall to the trap of degrading the schedule and allows adapting to the latest information available. As the literature review suggested, the schedule adaptation approach suffers from lack of time for calculation of assignments at the peak time when the complexity of assignment is the highest, showing that the system is not robust enough to handle demand peaks.

Therefore, in order to be able to handle any combination of immediate and advance booking requests, this approach proposes a real-time taxi dispatching, which in time intervals solves a static problem with the latest information in a constructive re-assignment framework. Re-assigning booking trips can leverage on the simultaneous assignment of a group of booking requests, but only given certain pre-requisites: (1) First, because the bookings and taxis are spatially and temporally bounded, only bookings with similar request times and locations can be re-assigned in order to improve system efficiency. (2) Second, the number of booking requests must not exceed the number of taxis. Otherwise, some taxis will be assigned not one booking but a chain of bookings. This is not assumed in this study, with respect to solution methods described later, which assume pairwise matching.

Overall approach

Buffered passenger requests at time intervals at the beginning of each decision epoch (visualised in Figure 15 by the diamonds in the first row) are assigned (inserted) to the schedule (represented by the bars in the third row) and passengers receive the response whether their request is accepted or rejected (visualized by the diamonds in the fourth row). The subsequent re-assignment process uses the time until the beginning of the next decision epoch to improve the performance of taxi operation (visualized by the bars in the second last row) and the updated schedules are published at the end of each decision epoch (visualized by the diamonds in the last row). During one decision epoch, the whole taxi schedule is re-assigned as described above. The same process then repeats in regular time intervals, as illustrated in Figure 15. The re-assignment requires some time. During re-assignment new requests cannot be accepted and are buffered. In order to minimize the delay of the initial reply to the passenger's request, this study divides the decision epoch into assignment (insertion) and re-assignment. (Kümmel et al., 2016c)

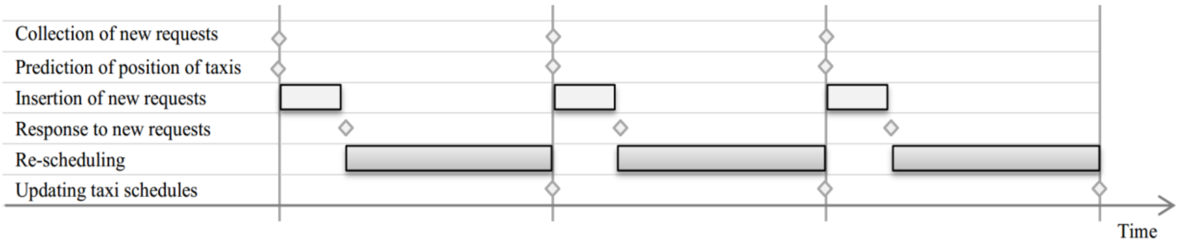


Figure 15 The assignment and re-assignment within decision epochs. (Kümmel et al., 2016c)

Assignment of new booking requests

The objective of the insertion procedure is to evaluate quickly whether the new booking request can be served and, if so, to insert it into the schedule of the best-fitting taxi. Hence, the insertion routine in effect asks each taxi, 'Are you able to pick up the booking request with the specified origin, destination and desired pick-up time interval? And if so, what does it mean for your schedule costs?'

If a new booking request can be added, the dispatching system inserts it in the schedule of the most suitable taxi. The criteria for insertion feasibility and selection of the best-fitting taxi are selected as follows. The feasibility of booking request assignment (as well as re-assignment) must be examined from the perspectives of both the passengers and the taxis. From the passenger's perspective, the insertion is feasible only if the booking request can be accommodated within the pick-up window, which is defined by the desired pick-up time and the time period, for which the passenger is willing to wait. From the taxi operator's perspective, the insertion is feasible only if the booking request does not violate the driver's ability to fulfil other already confirmed booking requests and shift times.

Figure 16 shows the model of insertion process at a glance. In a nutshell, each passenger (booking requests) asks through the dispatching system taxis: "Can you pick me up and, if yes, when?" in order to find out, the booking request are first tentatively inserted to the suitable place in the schedule of all taxis asked. If the taxi can pick up the requests in the time interval the passenger desires to be picked up while not violating pick-up of other requests later and without violating taxi shift, the requests can be picked up by a taxi and is reasonable to calculate also the estimated pick-up time and costs of pick-up. Then the taxi can respond first to the dispatching system accordingly.

The dispatching system collects the answers and evaluates them. The dispatching system may treat immediate and advance bookings differently. For immediate bookings, it may be more important to pick up the passenger quickly (before the competition), whereas with advance bookings the time aspect is less critical and costs can be emphasized. In both cases, the system selects the 'best' taxi to pick up the passenger.

If more taxis qualify as the best (which may happen often with advance bookings), additional measures are employed to select just one. If none of the taxis can pick up the passenger, the booking requests are rejected and the customer is advised to book again later.

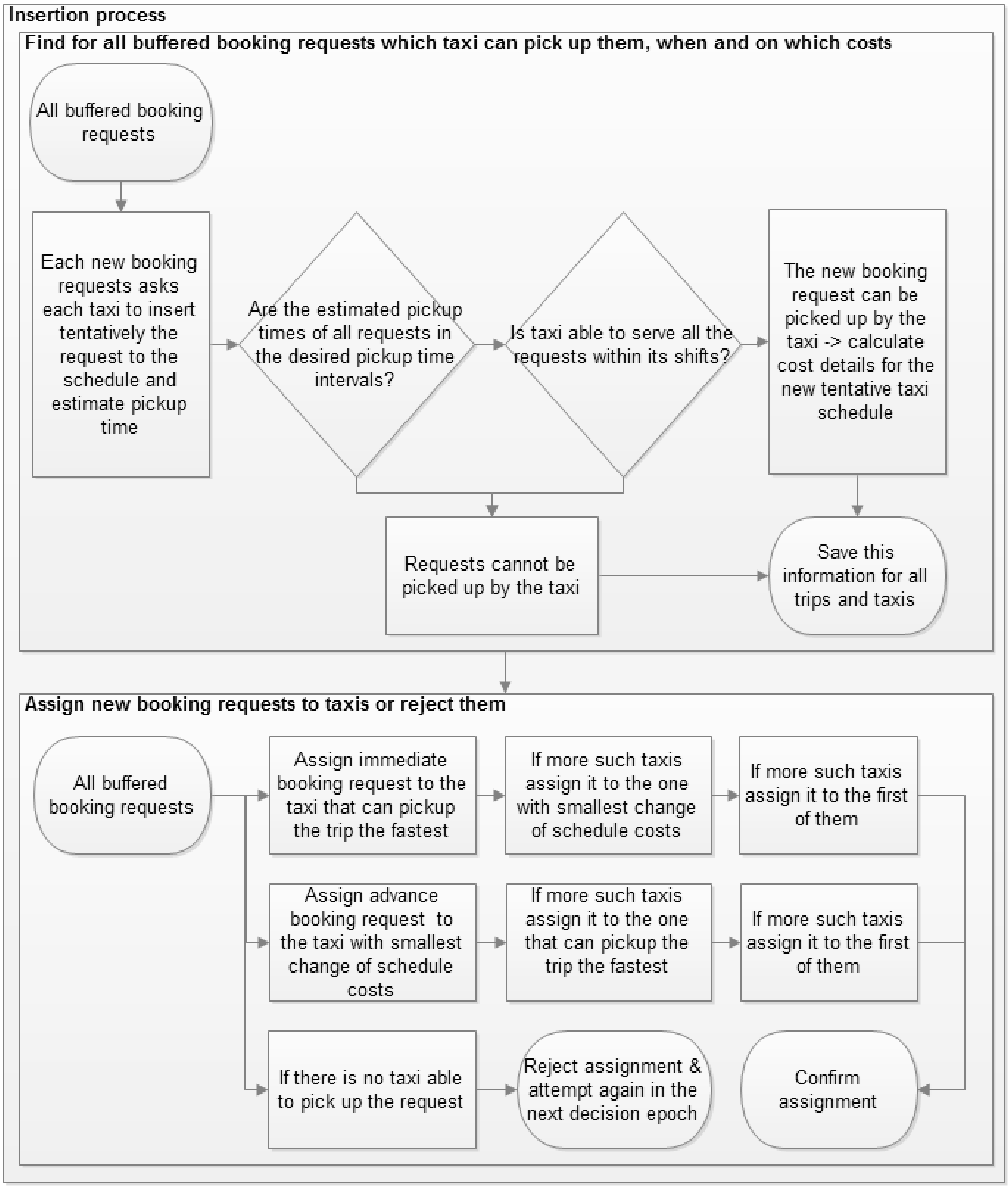


Figure 16 Insertion process in a nutshell

Unlike in Approach 1, the initial insertion of passenger booking requests to the schedule proceeds sequentially, because it is faster and all confirmed trips are re-assigned anyway immediately after that. Therefore, Figure 17 shows the classical sequential insertion to illustrate the insertion part of Approach 2. It is reasonable to base the assignments not on the position and status of the vehicles at the current time, but rather on their expected position and status at the end of decision epoch.

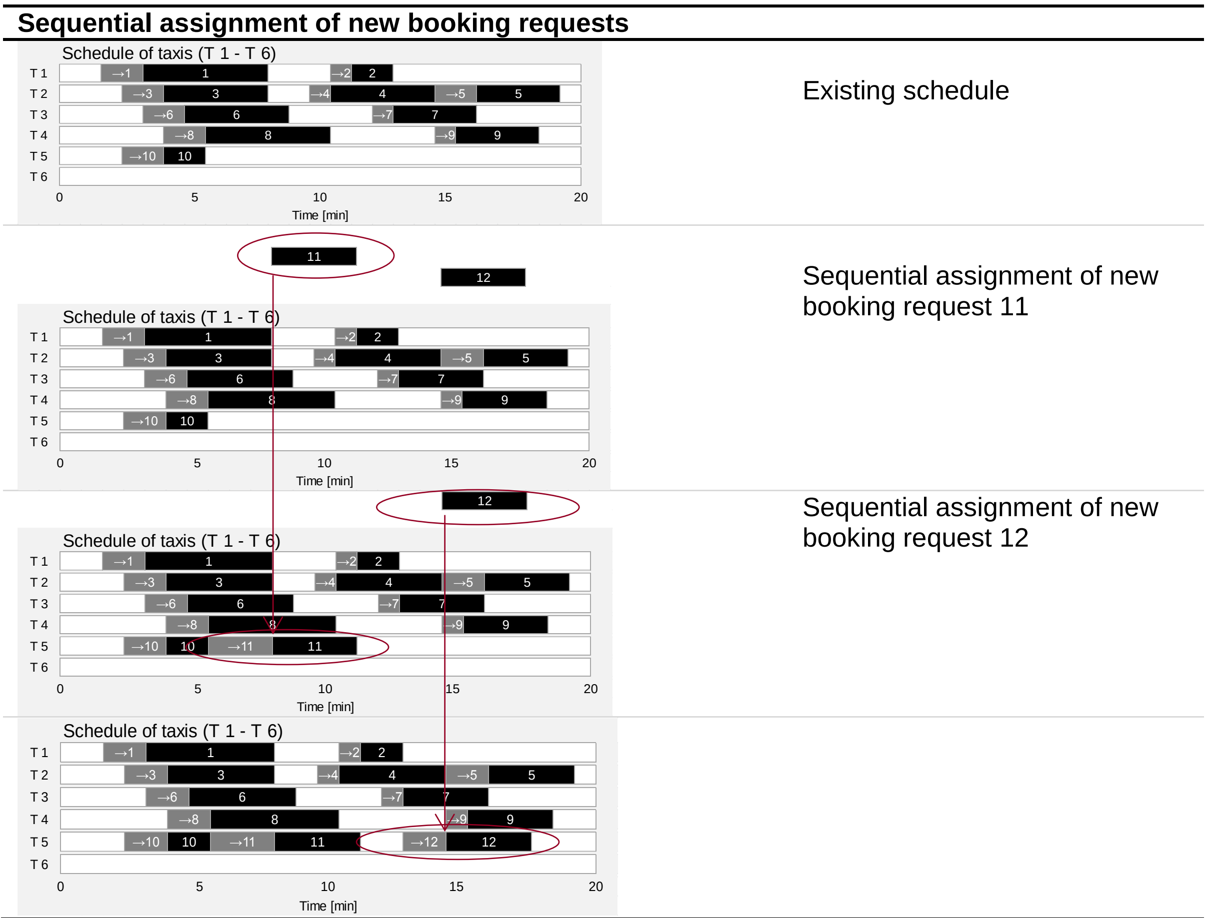


Figure 17 Sequential assignment of new booking requests

Re-assignment of already confirmed booking requests

The re-assignment, which happens just after insertion of new booking requests, proceeds as follows. It is reasonable to apply the group re-assignment sequentially in smaller time intervals one after another. These intervals start at the beginning of the scheduling horizon and at the end of the scheduling horizon. The beginning of the scheduling horizon is not the current time, but the current time plus the length of the decision epoch, because the re-assignment results will be known at this point. The end of the scheduling horizon depends on the advance booking time threshold, usually one day.

Figure 18 provides an illustration of the proposed re-assignment. In particular, it illustrates the process of removing of the confirmed booking requests and their simultaneous re-assignment. This figure depicts a schedule of six taxis T1 - T6 and their plan to pick up twelve confirmed booking requests 1 - 12. The re-assignment of confirmed booking requests begins with removing all confirmed booking requests from taxi schedules. Then the booking requests are assigned back to the schedule simultaneously in groups starting from the beginning of the schedule. Then the next group of booking requests are assigned the same way until the end of dispatching horizon, which is defined by the latest possible advance booking time allowed by the dispatching system – typically 24h. (Kümmel et al., 2016c)

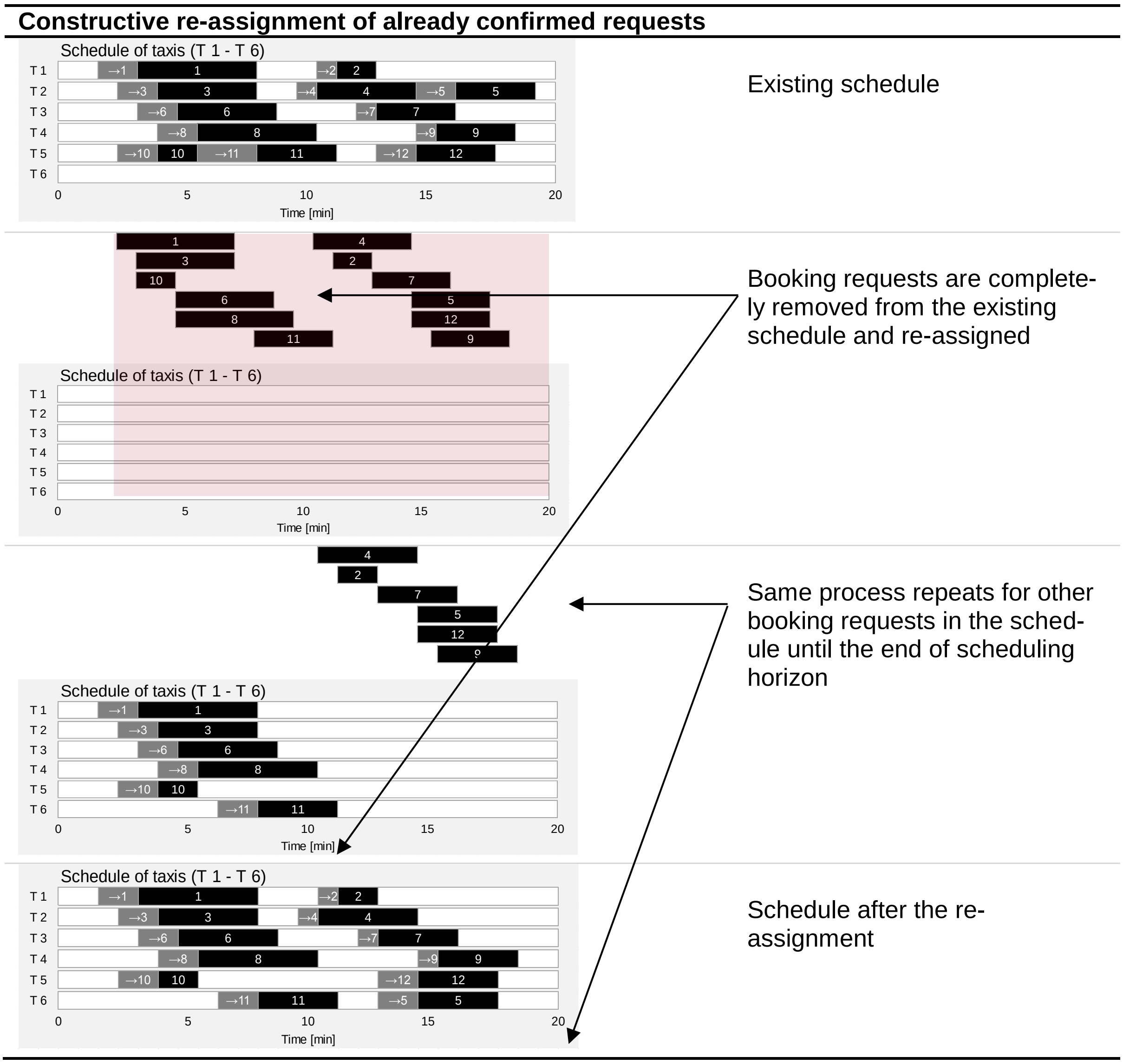


Figure 18 Constructive re-assignment of already confirmed requests (Kümmel et al., 2016c)

Figure 19 shows a model of the re-assignment. In the re-assignment (re-optimization or re-marriage) process, groups of passenger requests must be found first. The considerations for creating the group of trips and corresponding re-assignment time intervals follow.

First of all, each taxi searches in the schedule for booking requests, which can be re-assigned. This is true when the estimated 'on-call' time is larger than re-assignment time interval beginning. Only first such booking requests in the schedule of each taxi are considered. These trips must fulfil two criteria: (a) the estimated pick-up time the booking request must be in the future, (b) the estimated end point of 'on-call' time interval must be smaller than the re-assignment time end. If the request fulfils the criteria, it is added to the pool of trips to be re-assigned. The bilateral preferences of booking requests are calculated again, as they may have changed after insertion of new booking. Then a matrix of estimated pick-up times and costs for each request and taxi as well as the resulting preference matrix is created. Finally, the re-assignment can be calculated.

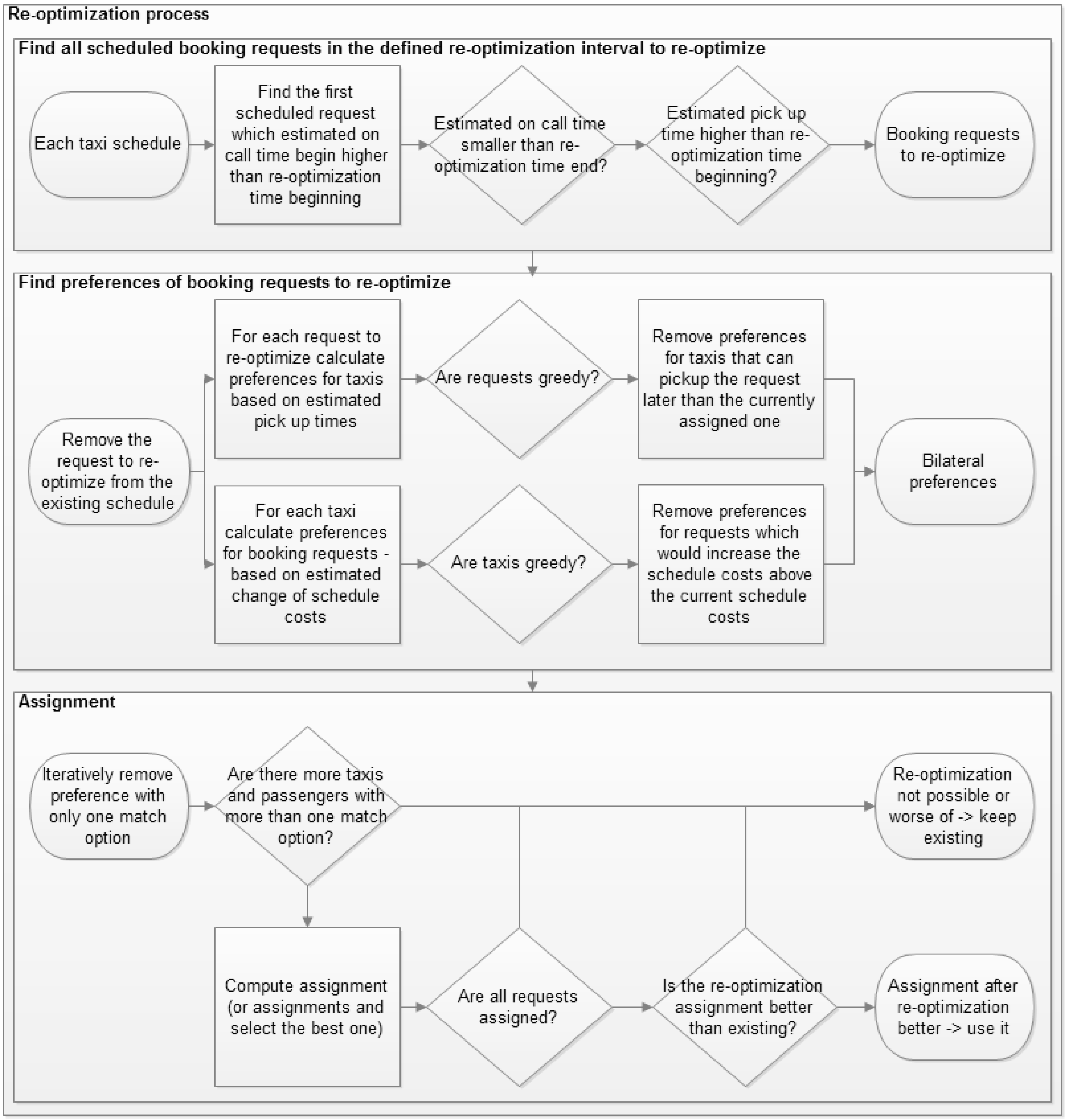


Figure 19 Model of re-assignment process

Constructive re-insertion suffers from the fact that occasionally not all booking requests can be re-assigned. Especially in high demand scenarios, the procedure, unfortunately, ends up not assigning some booking requests. In this case, the results of the re-assignment must be discarded and the schedule that existed after the insertion should be used instead. Therefore, it makes sense to iteratively remove all the assignments, which have only one available assignment option. This increases speed and more importantly prevents some infeasible assignments. If all booking requests can be re-assigned and the resulting assignment is better than the existing one, it is selected.

In a summary, the proposed methodology allows advance booking requests to be re-assigned in real-time. This is achieved by using the concept of constructive re-assignment of all confirmed requests.

3.1.5 Re-assignment of the timely nearest confirmed immediate bookings

The objective of Approach 3 is to investigate whether taxis should accept individually sub-optimal re-assignments for sake of system-wide performance.

This approach is an extension of the re-assignment Approach 2 by the greediness factor and by the ability to cope with driver shifts and breaks. Taxi drivers and passenger may be greedy; that means that they do not want to end up after re-marriage with someone, who is lower on the preference list than the current partner. That means that preferences, which are worse than existing (another taxi arriving later or another passenger with a higher pick-up costs) are not considered in the re-assignment process.

Unlike the Approach 2, the re-assignment considers only immediate bookings for two reasons. Firstly, the presence of shifts and breaks make scheduling challenging advance bookings. Secondly, the buffering interval should be very short, running every couple of seconds. Under these circumstances the proposed solution that re-assigns only the nearest group of already confirmed booking requests is very favourable.

In addition, unlike Approach 2, this allowed any change of the estimated pick-up time, so that a passenger with confirmed booking request may get re-assigned to another taxi, which may arrive later. Taxis also may be assigned to pick up more costly / less profitable passengers than already assigned one(s). The interesting question to ask is “Does it make sense to ask drivers during the re-marriage to not follow the natural greedy behaviour for the sake of system performance?”

Passengers do not accept any taxi, which will pick them up later than the currently assigned one (in that sense they accept only stable user-optimal propositions) but taxis may be assigned to pick up a passenger, which is not good for the taxis but good for the overall system performance.

Moreover, this experiment, unlike the previous two, is tested on real data, which contain taxi shifts and breaks. Previous experiments assumed continuous taxi availability. However, this experiment has to deal with breaks and shifts, limits the taxi availability and thereby adds constraints to the booking requests re-assignment process.

In the initial insertion phase, it is assumed that a new immediate booking is assigned to the taxi that can pick up the trip the fastest. If there is more than one taxi, then the taxi with the smallest change of schedule costs will pick up the booking. If there is more than one such taxis, the first taxi is assigned.

In the re-assignment phase, the bilateral preferences are constructed as follows. During the re-assignment, the passengers prefer the taxi that can pick them up the soonest. Taxis prefer to pick up passengers with the most favourable change of ‘on-call’ costs as compared to their current schedules. That means, not only the costs of the newly inserted trip, but the change of costs of all trips (including the newly inserted one) are considered.

There may be multiple stable solutions, all of which are stable but not equal in terms of performance. Therefore, it makes sense to introduce a simple measure that increases the number of assignments with the number of booking requests to re-assign. In the experiment this number is 2+n (where n is the number of booking request to re-assign). This is how these assignments are created: (1) Booking requests, who propose, are ordered according to least number of possible arrangements, so that the one with the least number of possible partners proposes first. This measure reduces the number of not assigned booking requests, because they have no matching option left. (2) Booking requests are ordered according to the desired pick-up time. (3) Booking requests are ordered randomly for n times. Naturally, it would be better to theoretically define, a method to compute the best stable results. However, with respect to the complexity of the problem and the number of modifications, this remains a future challenge.

In summary, Approach 3 extends the Approach 2 by adding the greediness factor and the ability to cope with drivers shifts and breaks. Therefore, it limits the re-assignment time period, which is not the whole schedule as in Approach 2, but only the one the timely nearest requests in the schedule. To counter-balance this, the re-assignment process is carried out more frequently than in Approach 2. Figure 20 displays an example of constructive re-assignment of the timely nearest group of requests. (Kümmel et al., 2016b)

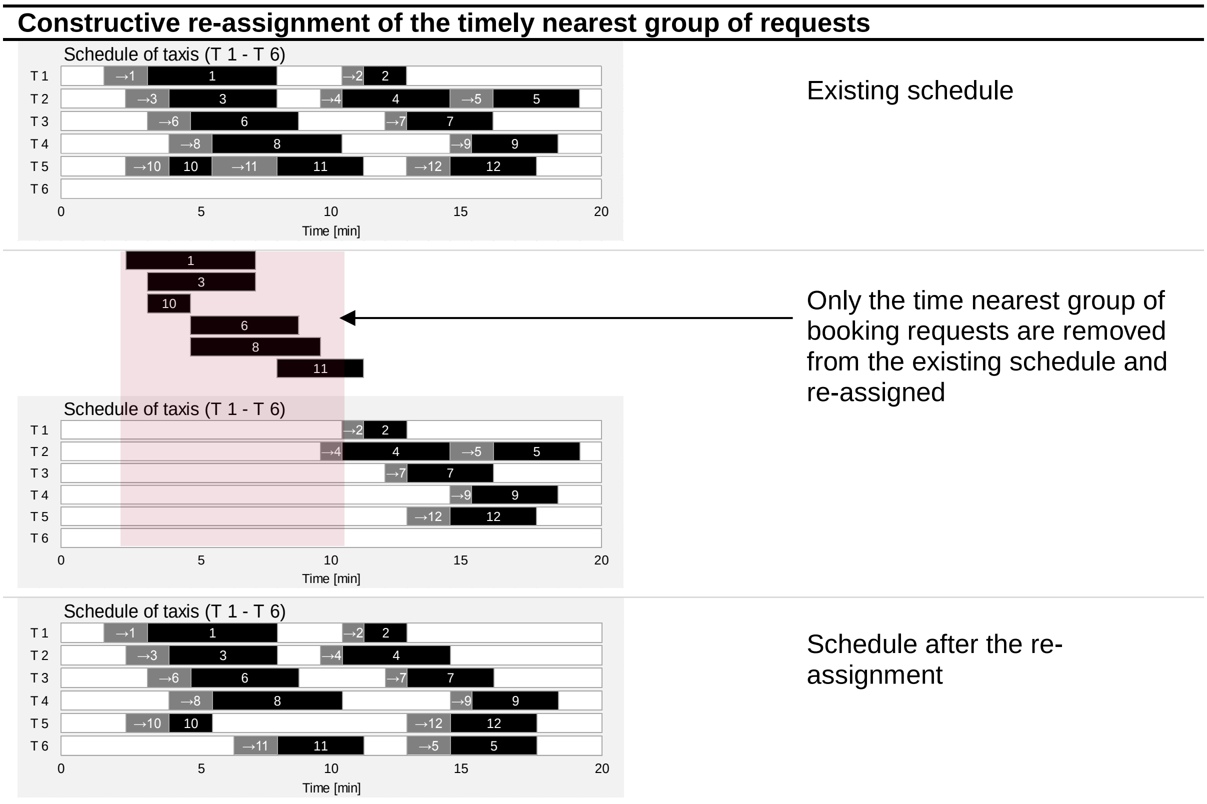


Figure 20 Constructive re-assignment of the timely nearest bookings (Kümmel et al., 2016b)

3.2 Implementation details

This chapter of the methodology investigates the implementation details of the proposed three approaches. First, it presents the detailed considerations of the process of assigning a passenger booking requests to the schedule, then it presents the original solution algorithm of the stable marriage problem and the necessary modifications required for its application for taxi matching. Furthermore, the details of implementation of the stable marriage algorithm; in particular, the preference making and order of proposal together with other considerations are proposed. Finally, a potential for group assignment is estimated on an example of three cities.

3.2.1 Details of assigning and re-assigning of passenger bookings

This chapter explains the assignment process in detail. The objective of the assignment procedure is to evaluate quickly whether new booking request(s) can be served and if so to insert it into the schedule of the best-fitting taxi. Hence, the assignment routine asks each taxi, ‘Are you able to pick up the booking request with the specified origin, destination and desired pick-up time interval? And if so, what does it mean for your schedule costs?’

The feasibility of booking request assignment must be examined from the perspectives of both the passengers and the taxis. From the passenger’s perspective, the assignment is feasible only if the booking request can be accommodated within the pick-up time window, which is defined by the desired pick-up time and the time period for which the passenger is willing to wait. From the taxi operator’s perspective, the assignment is feasible only if the booking request does not violate the driver’s ability to fulfil other already confirmed booking requests and while the taxi is on shift.

The assignment procedure is demonstrated in Figure 21, which depicts a schedule of five taxis and assignment of booking request 2. The ‘on-trip’ time (in black) is the travel time between origin and destination. The ‘on-call’ time to a booking request (in grey) is the travel time required to reach the passenger.

All taxis can serve booking request 2 except taxi 1, which cannot add this booking request without failing to pick up confirmed request 13 within the desired pick-up interval. These pick-up time intervals are visualised below each booking request by a petite light grey bar below. Because the desired pick-up times of booking requests 2 and 13 partially overlap, they are plotted one below the other. In contrast, taxi 4 can accept booking request 2 and still pick up the passenger 10 within the required time frame. (Kümmel et al., 2016c)

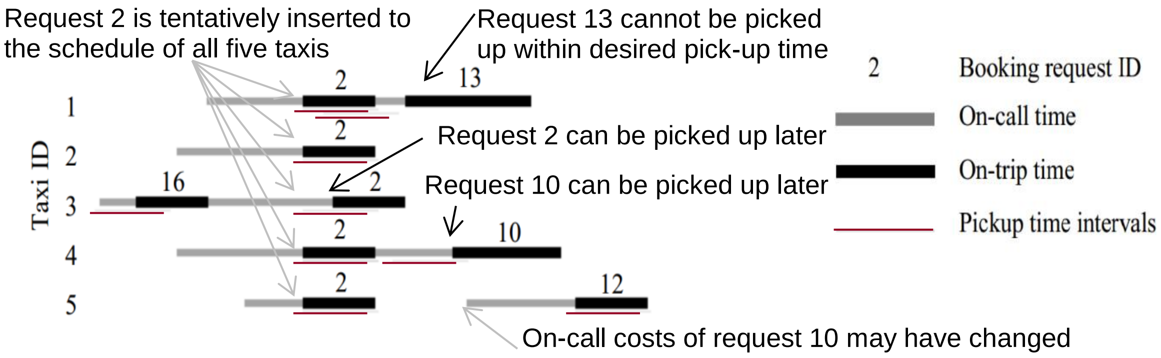


Figure 21 Example of assignment feasibility analysis for a booking request 2

3.2.2 Solution algorithm of the stable marriage problem

An assignment using the stable marriage algorithm, which matches two parties in a sense of Nash equilibrium, seems suitable for the stable taxi dispatching because it is a real-time simultaneous assignment method based on bilateral preferences. The algorithm aims to achieve a system-wide stable assignment of two parties at each decision epoch. Some authors such as Moore and Passino claim that a stable solution may be more useful than the close-to-optimal one in the presence of some disturbances such as communication delays or if short computation time harm the quality of assignment by traditional heuristic approaches. (Shamma, 2008)

The stable marriage algorithm is selected because of its ability to match two parties in polynomial time, models well the preferences of taxi drivers and passengers, works in competitive environment and guarantees stability thanks to Nash equilibrium. What makes the two sided matching problems different from the classical assignment problem (job-machine assignment) is that people on both sides (taxi drivers and passengers) care about the outcome of the assignment (Roth et al., 1993).

The solution algorithm of the stable marriage problem, proposed by Gale and Shapley (1962), has been applied in various situations, such as the matching of students and schools, doctors and hospitals and kidneys and patients. The objective of the problem is to match both parties in a stable fashion according to their preferences. The outcome (the assignment) guarantees that nobody has an incentive to change the existing matching (in the sense of Nash equilibrium). The algorithm facilitates the transition of competitive individualistic preferences into a stable situation.

In the original Gale-Shapley algorithm, men propose to women. Men and women are assumed to have preference lists about the other sex, as illustrated in Figure 22. Men propose to women (according to their preferences) and women response with either acceptance or rejection of men proposals (according to their preferences). The matching process runs in rounds as long as there is an unengaged man, who proposes to a woman.

Men	Men preference list	Match	Women	Women preference list
A	a, b, c, d		a	A, B, C, D
B	b, a, c, d		b	D, C, B, A
C	a, d, c, b		c	A, B, C, D
D	d, c, a, b		d	C, D, A, B

Figure 22 Stable marriage matching example (McVitie and Wilson, 1971) (Kümmel et al., 2016d)

Figure 22 provides an example of the matching of men and women. The matching begins with a randomly selected man 'man A', who proposes to the best woman in his preference list to whom he has not yet proposed 'woman a'. 'Woman a' has not yet received any proposal, so she accepts the proposal from 'man A'. Next, 'man B' proposes to 'woman b', and because she is not yet engaged she accepts. Further, 'man C' proposes to 'woman a', but she is engaged to 'man A'. She compares the proposal to her existing engagement and rejects 'man C' because she prefers 'man A' (he is higher on her preference list). Next, 'man D' proposes to 'woman d' and she accepts. Because there is still an unengaged 'man, C', in the next round he proposes to the next-favourite woman in his preference list to which he has not yet proposed 'woman d'. 'Woman d' is already engaged to 'man D', but she prefers 'man C' (he is higher on her preference list) and therefore swaps the engagement for the new proposal from 'man C'. This leaves 'man D' unengaged. He proposes to the most desirable woman on his preference list to whom he has not yet proposed 'woman c' and because she is not yet engaged she accepts. As there are no unengaged men left, the matching assignment ends. (Kümmel et al., 2016d)

The resulting assignment is visualised in Figure 22, with the paired men and women connected by lines. The original algorithm has two variants, in which either men or women propose first to the opposite sex. Both men- and women-centric variants may result in different assignment and are examined in the experiment. (Teo et al., 2001) investigated cheating to find out that chances to benefit from cheating are slim.

	General stable marriage problem	Taxi dispatch problem
Participants	Men and women	Available taxis and new passenger bookings
No. of participants	Equal (N*N)	Not necessarily equal (N*M)
Preference making	Arbitrary – based on personal taste of each man and woman	Passengers prefer a taxi that can pick them up the soonest Taxis prefer passenger requests with the least costs
Preferences complete	Yes, every participant of the matching has ordered preference list of all participants of opposite sex	No, some taxis may be too far away and therefore preference lists may be incomplete
Assignment produced	Once for given no of men and women and their preference lists	Every decision epoch for given no of available taxis and new passenger bookings and their preference lists

Table 15 The stable marriage and the taxi dispatch problem (Kümmel et al., 2016d)

Table 15 summarizes the key differences between the general stable marriage and the taxi dispatch problem. Figure 23 describes the assignment strategy based on the stable marriage algorithm. Just instead of (un)engaged women and men there are (not) assigned taxis and passengers. At the beginning the systems checks whether there are any 'vacant' taxis, which are able to pick up passenger(s). Then a randomly selected taxi proposes to the most preferred passenger (the same way as the unengaged man proposes to the best women in his list in the example above). The passenger proposes with one of the three possible answers: (1) if yet unengaged the passenger accepts, (2) if already engaged but receives a better offer, the passenger accepts the new offer and cancel the worse taxi offer and (3) if already engaged but receives worse offer to what the passenger already have, the passenger rejects it. A logical matrix keeps track of propositions and engagement statuses. If there are still unengaged taxis, who can pick up passengers, the assignment process continue in next round. Only when there are no more taxis that can pick up passengers the assignment process stops.

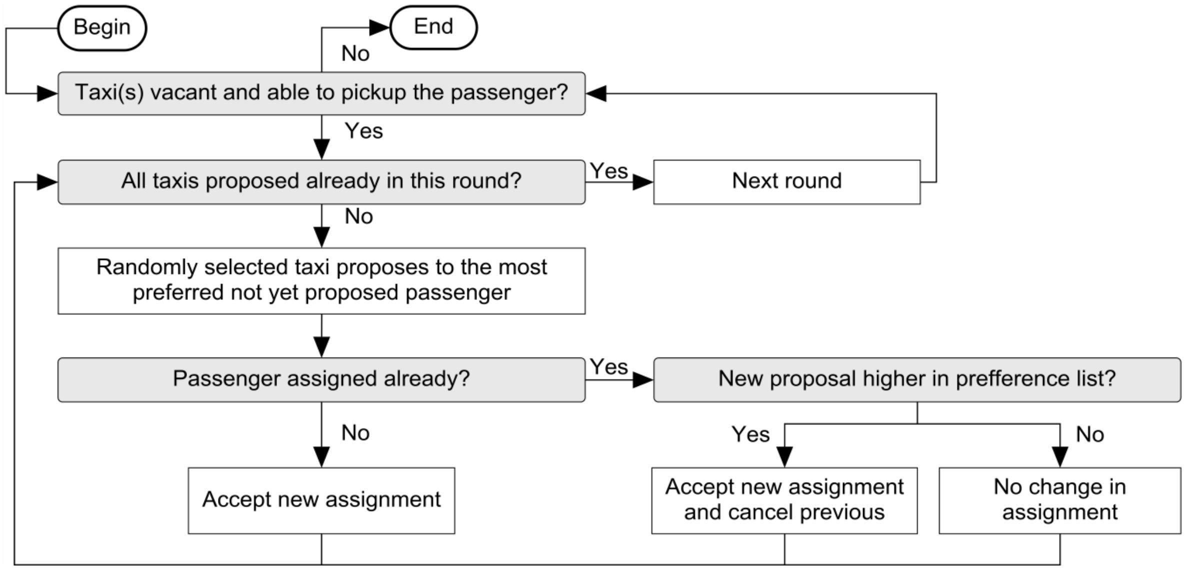


Figure 23 The assignment process based on the stable marriage algorithm

3.2.3 Preference making

The trips and taxis are assigned according to the preferences. The preferences of taxis and passengers are constituted in the following fashion. Passengers prefer a taxi that can pick them up first; taxi drivers, on the other hand, prefer booking requests with the least costs such as used by for example Bai et al. (2014). In contrast to the stable marriage algorithm, which assumed the same number of men and women having clear preferences without ties, the application for taxis and passengers requires a couple of modifications.

Limited willingness to wait and incomplete preference lists

It would be unrealistic to assume, as some studies do, that the passengers are willing to wait a taxi for an infinite time. Most likely, they will find another means of transportation. Therefore, it is reasonable to limit the passengers waiting time reasonably. Not every taxi can reach every passenger in the desired pick-up time and therefore the preferences cannot be created and the preference list becomes incomplete. Moreover, more taxis can pick the booking request up at the beginning of the desired pick-up window, about which the passenger is indifferent and the preference list becomes preference ties. These incomplete preference lists were theoretically described by Brito and Meseguer (2006) and formulated as an extension of the original stable marriage algorithm. (Kümmel et al., 2016c)

Preference ties

It may happen that two taxis can pick up the same trip at literary same time. How to organize the preferences then? The original stable marriage algorithm does not assume that for example for a women there will be two or more man equally attractive. Fortunately, the extensions of this algorithm with ties have already developed by Brito and Meseguer (2006).

Order of proposals

Moreover, in the very detail of application it also makes a difference, in which order either taxis or passengers propose. The original stable marriage assumes a random order. In the taxi centric version the first taxi that proposes is the one with least costs difference out of all taxis. In the passenger centric version, the first passenger to propose is the one, who needs least time out of all passengers to be picked up. Admittedly, there is a demand for a study that would scrutinize variations of the proposition order together with other variations of the stable marriage algorithm in order to further fine tune the performance.

Which party proposes first

The algorithm allows two variants, either taxis or passengers propose first to the opposite party. Each variant has a stable solution that may favour more either taxis or passengers depending on whom makes the first move.

Motion prediction

The decision epoch concept delays informing the drivers and passengers about their assignment. Therefore, it is reasonable to base the scheduling calculations not on the positions and statuses of the vehicles at the current time but rather on their expected positions and statuses at the end of a regular time interval. Since the horizon of the prediction is very short (equal to the length of a decision epoch), a linear estimation of taxi positions and statuses can be used. A detailed study on the effect of prediction on the performance especially in highly fluctuating and stochastic scenario would be desirable.

Veto

Some studies give the taxi drives a choice of accepting or rejecting the assignment such as (Seow et al., 2008), while some not. This study considers the drivers as agents acting on behalf of the central dispatching system and therefore does not allow drivers to veto the assignments. To more detailed extend, the acceptance and compliance with the assignments and, in particular, re-assignments are studied in Approach 3.

Priority

Some studies assign some customers higher priority than to others such as (Bullo et al., 2011). Some passengers may bring more revenue than others and therefore may be prioritized in the dispatching. While, this is a practical business arrangement, it is not considered yet in this study, because it would be very difficult to model this situation based on real data. Even if the dispatching companies would use such a priority scheme, it is difficult to imagine that they would be willing to disclose any details.

Illustration of a greedy matching

Aren't the preference ranks the same for taxi and passengers? This is true when there is just one passenger and one taxi; however, with increasing numbers of passengers and taxis, this is not necessarily true. The following paragraph and figure demonstrates why.

Figure 24 depicts a situation with two passengers and two taxis and their preferences. Both passengers prefer the taxi that can pick up them the soonest, which is in this case the one in between them. The taxis prefer to pick up the closest passenger, which is the one on the top right corner. This proves that the preference ranks must not necessarily be the same for taxi and passengers.

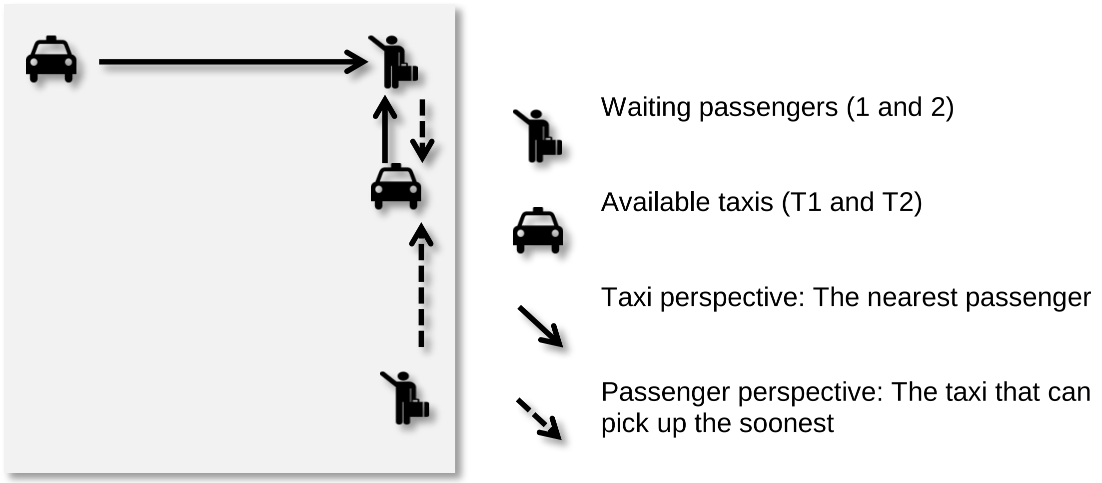


Figure 24 Bilateral preferences do not necessarily constitute the same preference ranks

Potential for simultaneous (re-)assignment in selected three cities

The critical mass of booking requests is different in every city. The following paragraphs attempts to roughly quantify the critical mass and therefore suitability of a re-assignment strategy for Munich, San Francisco and Singapore. There is no known guideline for quantifying the suitability and the proposed method uses many simplifications, nonetheless it serves as a basic indication of potential for re-assigning in a city.

The critical mass is expressed as a ratio of a total theoretical area, from which the booking requests can be picked up over the area of a particular city. The higher the ratio the higher are the chances that there may be some booking requests that can be re-assigned to be served by other taxis and thereby the performance improved. Note that this number is a subject of many simplifications and assumptions, however, has the capacity to indicate a potential for improvement in various cities and their taxi passenger demands.

A coverage area of one request is an area from which taxis can pick up a booking request. The area can be expressed by an area of a circle A, where the radius r is a distance, which a taxi can cover within the passenger is willing to wait. The radius r is higher with a higher velocity v and a higher willingness to wait (WW). The total theoretical coverage area of all booking requests (TCA) depends on the area size (A) and the number of requests (N) in a particular time. Then the ratio of the potential is the total theoretical coverage area divided the area of a city (AC).

$$R = \frac{TCA}{AC}, TCA = A * N, A = \pi r^2, \qquad \mathrm{r} = \mathrm{v} * \mathrm{WW} \tag{9}$$

In this example, it is assumed that in Munich, San Francisco and Singapore taxis move freely with constant speed of 36 km/h in a link-less area. Passengers are willing to wait for a taxi for 1000 s. Singapore has on average 1.02 million trips per day, which corresponds to 708 requests per minute, out of which one half (354 trips) is booked. A standard deviation in the booking demand of 177 trips is assumed. San Francisco has assumedly on average 50000 trips per day, which corresponds to 35 trips per minute, out of which one half (17 trips) is booked. Standard deviation of 16 is assumed. Munich has on average 46 575 trips per day, which corresponds to 32 trips per minute, out of which one half (16 trips) is booked. Standard deviation of 16 is assumed. The respective areas of Singapore, San Francisco and Munich are assumed to be 718, 122 km^2 and 310 km^2. (Fang et al., 2007), (Land Transport Authority in Singapore, 2014) (Landratsamt München, 2016)

As a result, given the assumptions made Figure 25 estimates the limits of applicability of simultaneous assignment of taxis to booking requests. In case of Munich the average booking rate of bookings covers approximately 8 to 24 fold the area of Munich. This creates reasonable chances that some trips origin are close enough in time and space so that they can be re-assigned. For San Francisco the chances are slightly higher thanks to smaller area, so that the bookings cover an area 13 to 34 times the San Francisco area. In case of Singapore, the average daily booking demand creates an area cover, which is approximately 78 to 232 times higher than area of Singapore. This means that there are substantially higher chances of exchanging some bookings even during the low demand. On the other hand, as in all cities there are many taxi operators and dispatcher, the actual potential is lower – corresponding to supply and, in particular, demand size. Nonetheless all considered cities are assumed to have enough potential for re-assigning.

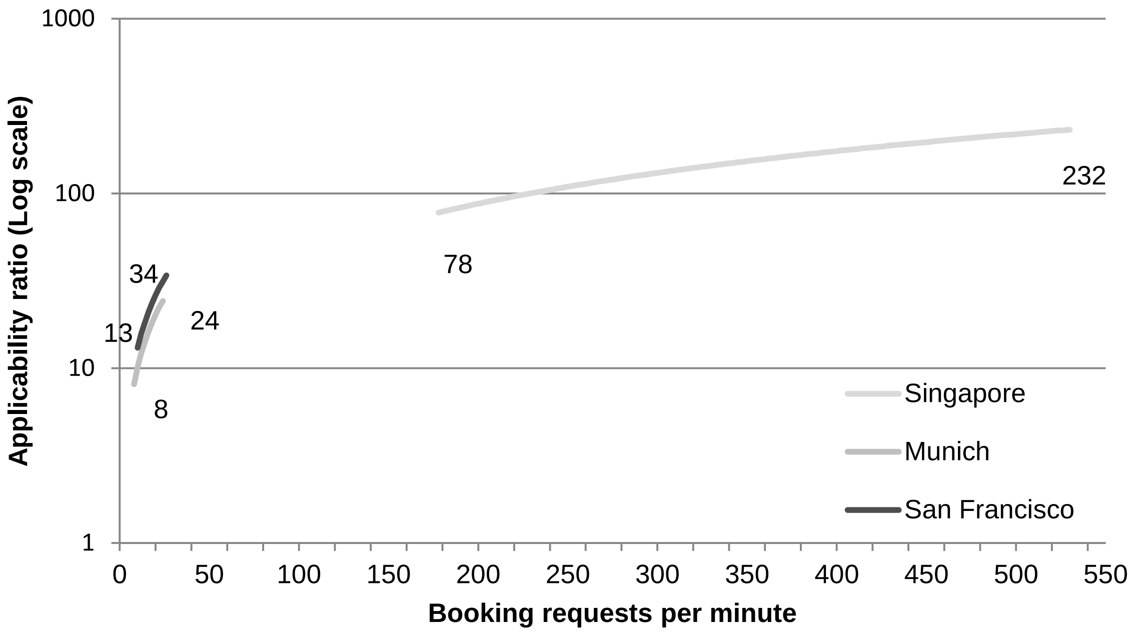


Figure 25 Applicability ratios for taxis in Munich, San Francisco and Singapore

3.3 Towards full automation of dispatching and operation

Automated taxi vehicle fleets will reshape the urban public transport systems. This chapter focuses on challenges and opportunities of the transition from conventional taxis with drivers to automated vehicles and fully autonomous taxi fleet management. The objective is to propose a theoretical framework for the fully autonomous fleet management. However, it is not the intention of this work to come up with some definitive one solution for the scheduling as such, as this is beyond the possibilities of this thesis and would require a coordinated effort of many teams for many more years. This study specially considers high demand automated taxi systems in megacities, in which the demand for taxis often exceeds the supply.

This study aims to not consider only replacing conventional taxi vehicles with autonomous ones but on creating new ecosystems or models of automated operation. In particular, removing taxi drivers and their decision making is an impulse to comprehensively reconsider the whole taxi system and all its operations and to lay foundations for the future connected, data-rich and almost fully automated system. Hence, which are the activities in the taxi systems controlled by humans that need to and can be replaced?

Taxi drivers, apart from driving, constitute a substantial decision making body of the taxi system. They are typically responsible for service, maintenance and cleaning and refuelling / recharging of the vehicle. Moreover, depending on the particular business arrangement, taxi drivers define or co-define their working hours and breaks, pre-allocate their taxis in the expectation of future demand, search for passengers by cruising on streets or waiting at taxi stands or stations and are accepting or rejecting assigned passenger booking requests. In some taxi systems, the drivers are more independent than in others in making the above mentioned decisions.

Fleet managers, dispatchers and dispatching systems influence fleet operation either strategically or operationally. For example, on the strategical level, the general business strategy, the size of the fleet and fare structure needs to be determined. On the operation level, the dispatching strategy how to assign booking request needs to be set. Coordinating taxis to serve locations with a higher expected taxi demand than supply results in a meaningful competitive advantage. Figure 26 summarizes the above mentioned activities and their corresponding current decision makers.

Activities / Decisions	**Who decides?** Taxi drivers	Fleet manager
Driving	✓	
Service, maintenance and cleaning	✓	
Refueling / recharging	✓	
Accept / reject passengers	✓	✓
Pre-allocation / passenger-search	✓	✓
Defining working hours	✓	✓
Dispatch taxis		✓
Estimations of taxi supply-demand		✓
Define fare structure		✓
Estimation of needed fleet size		✓

Figure 26 Taxi activities and decisions owners that can be replaced

This study presents a framework for managing automated vehicle fleets autonomously. In particular, it uses a metaphorical description of a family, which helps to decompose the complexity of multifaceted taxi operation and provides a framework for the automated vehicle fleet management. This metaphor models all activities and decisions in a transparent and sizeable manner. Moreover, it allows for variations in terms of centralized and decentralized decision making. The proposed family members are:

- Mother – strategic manager
- Children – automated vehicles
- Father – provider of physical services

The following paragraphs elaborate the responsibilities of the family members.

Mother – Strategic manager

The mother represents the fleet strategic manager, dispatcher and adviser. She is responsible for the family's economy and customer satisfaction. In general, the mother can act either more authoritatively (exercise centralized control over every aspect of the children's behaviour) or more liberally (set targets and provide recommendations instead of commands). She may even change her way of acting, in the dependence on the market conditions and competitors' behaviour. These are the assumed main three roles of the mother:

In particular, the mother defines fleet priorities and strategy business goals. The mother may increase or decrease the number of children in the fleet or adjust fares in response to the market development. The mother may also define the vehicle availability during a day.

The mother may manage the scheduling and re-scheduling of the trips, that is to ask the children to serve a particular booking request or exchange it with other child. In this sense, the mother may replace the traditional dispatching system.

The mother may also ask the children to pre-allocate. The vehicle pre-allocation may help her to further improve the service, such as to reduce passenger waiting time, which, however, increases costs. She may change the priorities of the company depending on the market situation (competitors' strategy and customer expectations). Moreover, the mother may control or at least set rules for using the physical services, as described later.

Moreover, the mother collects statistics of the past passenger demand and vehicle data and predicts the future demand and supply balance. She provides this big picture and transforms the information to actionable insights for the children.

The mother is receiving passenger booking requests. She communicates with the customers and the family members. The mother at the last instance collects a share or whole revenue from fares.

Children – Automated vehicles

The children in the proposed framework represent the individual autonomous vehicles. The children like to play (drive passengers) and intend to make the mother happy (by fulfilling her targets). A way how to make the mother happy is, for example, to decrease fleet costs.

If the mother is liberal, taxis may decide about the assignment of passenger booking requests. They may even negotiate and re-assign them in order to reduce costs of serving them. Fleet costs can be reduced, for example, by swapping some of the scheduled passenger trips such as indicated in Figure 13. In order to do so, the children are able to estimate the costs and revenues of their scheduled trips. The costs and revenues are based on distance, time or an event. For example, the fares are either based on distance or time and may contain also some one time surcharges (such as airport surcharge). The costs arise while driving based on distance (constituted for example by tear and wear), based on time (constituted for example by a prorated vehicle purchase or lease price) or based on a one-time event (such as expenses for energy recharging or cleaning). The children communicate in order to further improve their schedules because new passenger booking requests may degrade the schedule quality (as Figure 13 shows). Therefore, the children negotiate among themselves in order to find a better assignment of the passenger trips as well as the time slots for charging, service and maintenance.

If the mother is authoritative, children just provide information to the mother, who commands them. Both options of assigning the booking requests are possible. The advantage of the centralized authoritative control is decision making with the complete information. The advantage of decentralized liberal control is expected faster computational time of more details leading to more robust assignments.

Children collect a lot of data from their actions and from sensing the environment. They may partly process these data and send them to the mother. The data can be for example used for improving estimation precision of supply-demand balance and for real-time routing of other traffic participants.

The children communicate primarily with the mother and among each other about booking requests. They also communicate with the father about necessary physical activities such as recharging, refuelling cleaning, service and maintenance.

Father – Provider of physical services

The father provides the family with all necessary physical services such as recharging, refuelling cleaning, service, maintenance. In essence, the father is a booking platform for the children, in which they can schedule the required activity. As mentioned in the introduction, as this study takes a special consideration of high demand taxi systems, it justifies the existence of the father by assuming the constrained capacity of physical services and a need to manage their utilization.

The father requires the children to book the necessary services. The father also charges the children for using the services. As the demand for these services may fluctuate, the father requires the children to reserve these services in advance in order to avoid waiting for a free slot. The father may use static or dynamic pricing, depending on the particular fluctuation of the demand. The service reservations creating a facility schedule, however, may be changed later.

The father may gather statistics and provide predictions about the expected availability of the service spots and the expected price. The father is also communicating with the rest of the family. He may also charge the mother, for providing the predictions of availability. In the practice, automated vehicles, infrastructure and service may communicate via application programming interfaces. These interfaces enable separate business entities to communicate through one platform and thereby allow the transparent use of an infrastructure.

Figure 27 summarizes the possible shift of responsibilities in the proposed automated ecosystem based on fully automated decisions. The grey bars represent the possible carrier of the responsibility or the influencer. Depending on the particular setting, there are options possible.

Activities / Decisions	**Who decides now?**		**Who overtakes responsibility in the future?**		
	Taxi drivers	Fleet manager	Children – Automated vehicle	The mother – Strategic manager	The father – Service provider
Driving	✓		✓		
Service, maintenance and cleaning	✓		✓	✓	✓
Refueling / recharging	✓		✓	✓	✓
Accept / reject passengers	✓	✓	✓	✓	
Pre-allocation / passenger-search	✓	✓	✓	✓	
Defining working hours	✓	✓	✓	✓	
Dispatch taxis		✓	✓	✓	
Estimations of taxi supply-demand		✓		✓	
Define fare structure		✓		✓	
Estimation of needed fleet size		✓		✓	

Figure 27 Responsibilities in the proposed automated ecosystem

Principal processes and communication streams

Figure 28 summarizes the whole framework. The detailed description of the process and scheduling is described in the next paragraphs. Figure 28 illustrates the main processes and communication streams of the framework among the family members. When a passenger requests a taxi, the mother receives this request and asks the children whether they can pick up this request. The children reply in what time and at which costs they can serve the request and the particular booking request is assigned (either as a result of centralized decision by the mother or children decentralized negotiation) to the particular vehicle based on the current priorities. Then, the mother replies to the passenger by confirming the booking request or by rejecting it (if there is no more taxi available, which can serve the booking request). The next section details briefly the aspects of the family based scheduling framework proposed and described above.

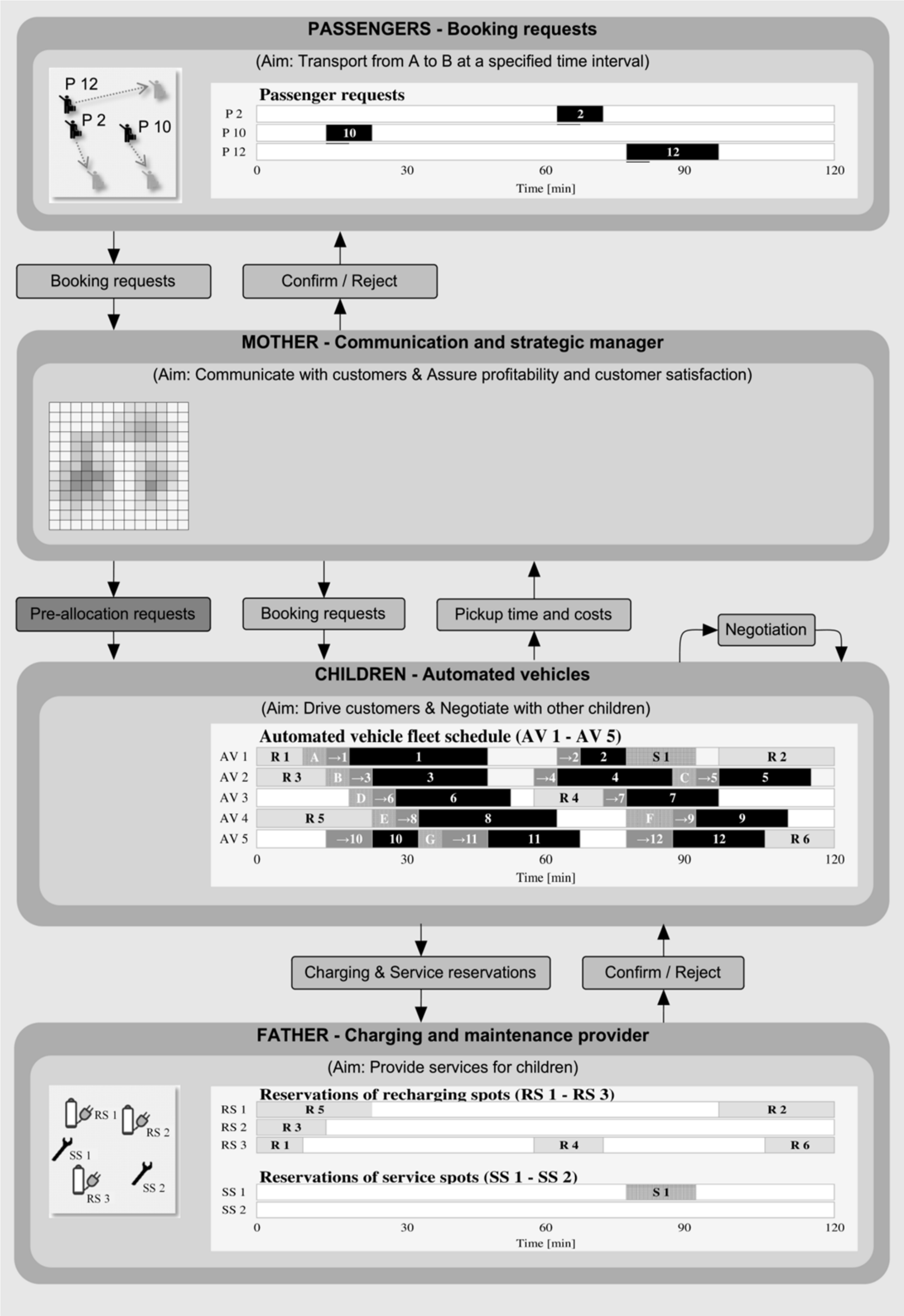


Figure 28 Family based framework for automated taxi fleets

Figure 29 represents a simplified expected taxi-supply demand balance. Based on this estimation, the mother may ask some children to pre-allocate to the favourable positions. The favourable positions are either expressed by the balance of supply and demand or by the profitability or likelihood of pick-up. Moreover, these may not be based on areas, but roads instead.

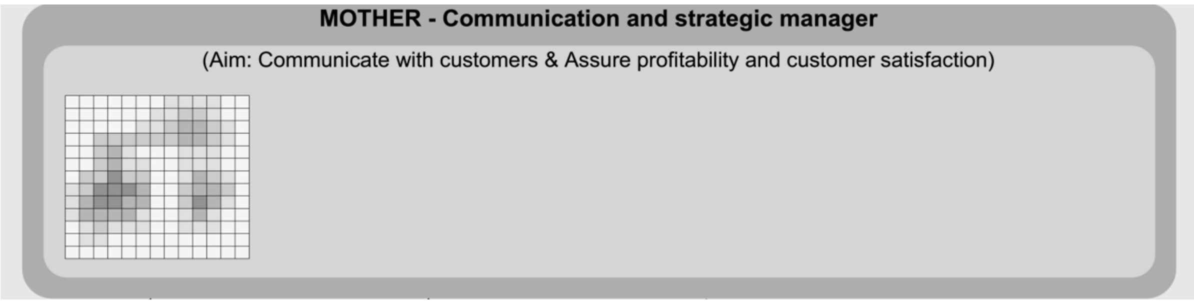


Figure 29 Prediction of expected taxi-supply demand balance

Service infrastructure and their booking is illustrated in Figure 30. On this example only recharging and service spots are shown, however, also the maintenance, cleaning and other services may be booked the same way. It is assumed that each activity has certain duration and can be done on certain location. That’s very favourable, because it can be represented on the schedule in a very similar fashion as trips with passengers.

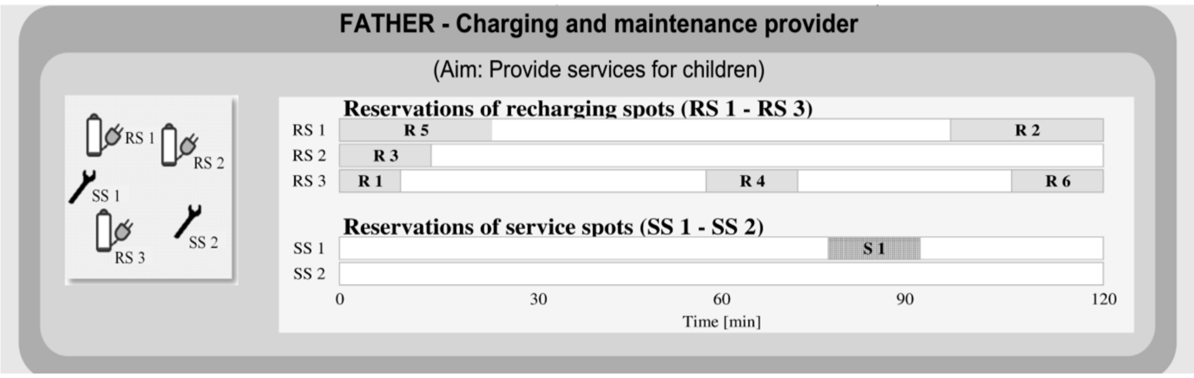


Figure 30 Service infrastructure and their booking

Figure 31 illustrates the proposed integrated scheduling framework for automated vehicle fleets of five automated taxi vehicles (AV 1 – AV 5) and twelve passengers (1 – 12). Each trip with passenger (1 – 12) is preceded by a trip to passenger (so called ‘on-call’ trip), in which the taxi drives to the origin of a passenger trip (→1 – →12). The taxi status may either be ‘occupied’ (on a trip with a passenger), ‘on-call’ (on a trip to a passenger) or ‘vacant’. The bottom part of Figure 31 depicts the cumulative taxi fleet status. It contains also the pre-allocation (A – G), recharging (R 1 – R 6) and service trip (S 1). Reservations of charging and service slots are illustrated at the bottom of Figure 31 with three recharging spots (RS 1 – RS 3) and two service spots (SS 1 – SS 2).

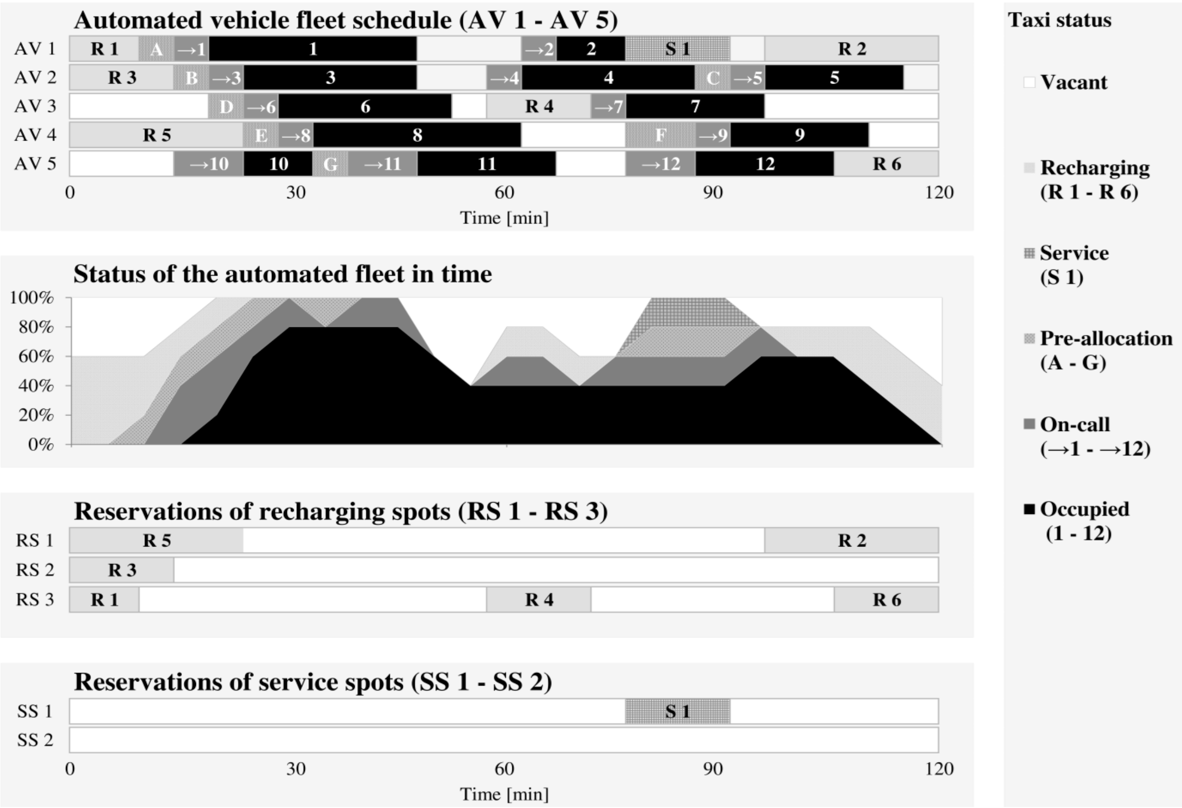


Figure 31 Proposed integrated scheduling framework for automated vehicle fleets

In the reality, there may be more mothers (fleet managers) and fathers (infrastructure and service providers). Each mother represents a unique company. The ultimate know-how of a mother to win over the competition is the right balance of fleet costs and customer satisfaction. Each father represents an unique company, which provides the recharging, cleaning and other services.

Many services may actually disappear or merge with other services. For example, there will be no difference between operations of automated taxis and automated car-sharing. For example in case of car-sharing, the vehicles also need to be pre-allocated, recharged and serviced. The trips with shared automated cars can be pre-booked and the car itself can come to the required origin, so that the customers do not have to search the vehicles in stations or in streets. Moreover, automated vehicles may cross-functionally deliver urgent urban post and parcels, food or medical supplies. There does not need to be a dedicated fleet for one purpose. On the contrary, sharing would improve utilization.

This approach, in particular, decomposes the arising complexity of the scheduling, which will take over decision making previously done by human drivers. On one hand, it allows decentralized negotiations among the vehicles that search for improvements of their schedules; on the other hand, it allows centralized controlling of the fleet objectives and priorities. Furthermore, it proposes a transparent framework for communication with infrastructure and service providers. The framework allows scheduling of various trips and tasks and, therefore, can be expanded beyond automated taxi scheduling.

3.4 Simulation model

The impact of the two proposed approaches to taxi dispatching and scheduling needs to be evaluated in a simulation model. Despite many models being developed in taxi research literature, at the moment of making a decision about the evaluation tool, there is no suitable simulation model available. Therefore, the model is developed 'in-house'. Table 16 summarized key assumptions used for the developed simulation model. The simulation model precisely emulates the dispatching strategies and allows any combination of immediate and advance booking requests, which are dynamically coming to the system. On the other hand, it consciously compromises some details in representing the environment – such as on the underlying network and traffic.

This simulation model, despite these simplifications, should, however, serve the purpose / be sufficient in the first exploration / evaluation of the proposed methodologies. More detailed simulation models, such as for example by MatSim with recently developed plugin for vehicle routing, should be used to validate the approach. (Maciejewski and Bischoff, 2015) Table 16 highlights in light grey by the key assumptions, which of the simplified or advanced ones are considered.

Categories	Abstracted / Simplified	Real / Advanced
Booking requests	Immediate (DOD=100%)	Combination of immediate and advance (DODϵ<0%,100%>)
Booking requests known?	Yes	No
Computation	Offline / static	Online / dynamic / real-time
Network representation	Euclidean space / simplified artificial network	Real network with full details
Traveling speed	Constant / historical distribution	Real-time
Estimation of traveling speed	Based on constant speed	Stochastic estimation
The taxi fleet size	Constant	Varying
Estimation of booking requests	Not used	Stochastic estimation
Return to depot after end of shift	Not required	Required
Non-revenue generating tasks (charging, refuelling, servicing)	Not considered	Considered
Learning effect of the driver	Not considered	Considered
Breakdowns of vehicles	Not considered	Considered

Table 16 Assumptions for the simulation model

The simulation model is developed by the author in MATLAB and its features are explained below and illustrated in Figure 32. The simulation model is organized as a discrete event-based simulation. The model schedules events (tasks) and executes them at predetermined points in time. Some events are known and scheduled in advance and others are scheduled only during runtime.

During the initialization phase, the initial positions of taxis, the origins and destinations of booking requests, booking times and desired pick-up intervals are generated according to the specified number of taxis, passenger booking requests and a specified distribution.

These events are scheduled to be executed at the predetermined time. All the key processes are visualized in Figure 32.

During runtime, new events are scheduled and scheduled events are executed. For example, if a new passenger request comes in (the scheduled event 'New booking request' is executed at a given time), the status of the booking request is changed to 'booked'. The event 'Taxi schedule calculation' represents the taxi dispatching, in which the new booking requests are either confirmed or rejected. Once a booking request has been confirmed, the following events, 'Passenger begins waiting' and event 'Dispatch taxi to trip' are scheduled. Once the event 'Dispatch taxi to trip' has been executed, the status of the taxi changes to 'on-call' and the event 'Taxi pick-up passenger' is scheduled. Upon execution of a 'Taxi pick-up passenger' event, the status of the respective taxi changes to 'occupied' and the respective passenger status changes to 'on-trip'. Similarly, upon execution of a 'Taxi drop-off passenger' event, the status of taxi changes to 'vacant' and the passenger's status changes to 'delivered' (or 'served'). The simulation ends when there are no more events waiting to be executed.

Booking requests are not known to the model until they are made and the decision as of which taxi to assign to which passenger is computed in real-time. The network is Euclidean space, in which taxis can move freely and pick-up and drop-off possible anywhere within this space. The travelling velocity is constant. The physical network, links, turn restrictions, variable travel times and many other features are not considered. While these assumptions do not reflect the reality, they are suitable for drawing initial conclusions and can be extended in the future. (Kümmel et al., 2016d)

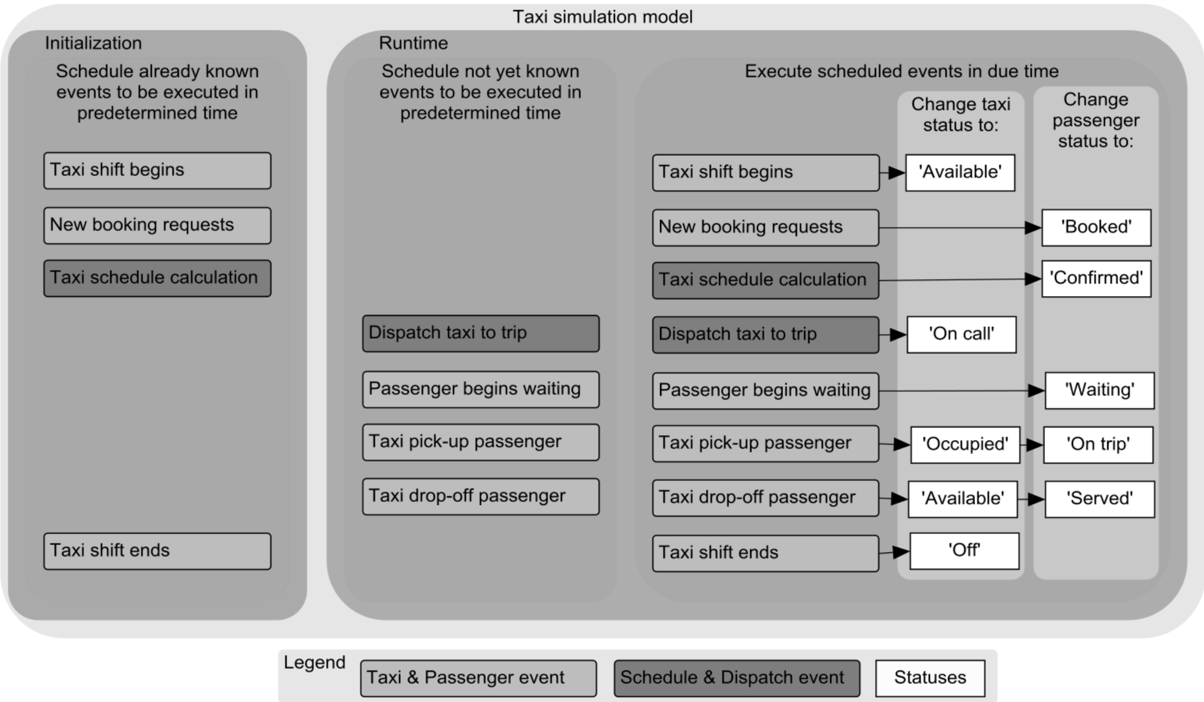
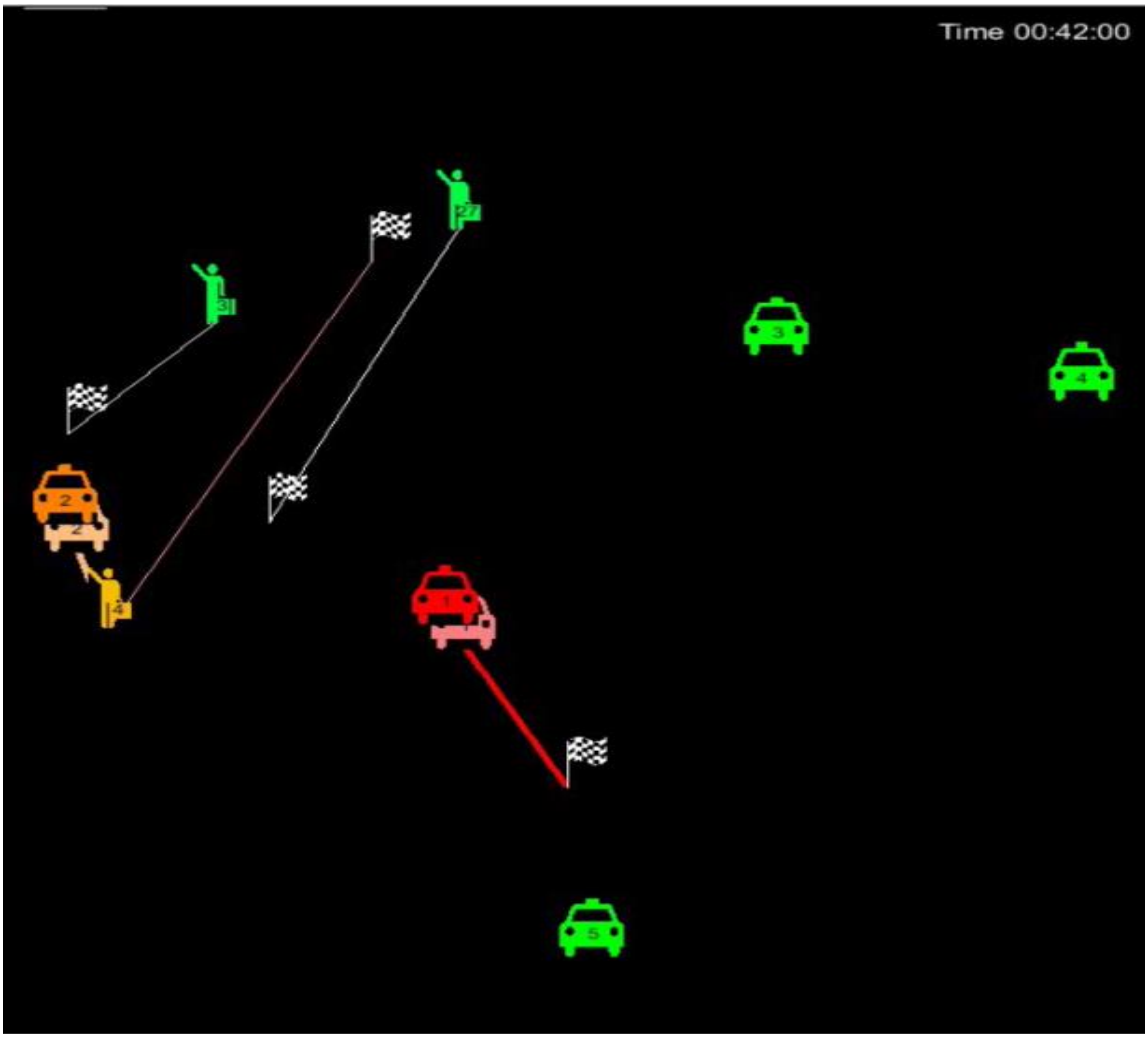


Figure 32 Taxi simulation model features and structure (Kümmel et al., 2016d)

The screenshot of the taxi simulation model in Figure 33 illustrates the spatial dimension of the taxi dispatching problem and Figure 34 the temporal dimension in a form of taxi schedules. The upper schedule is the schedule after insertion of new booking requests. The middle schedule is the schedule after the re-assignment process. The bottom schedule is the actually accepted schedule.

Spatial dimension representing a city



Legend

Taxis ‘Vacant’, ‘On-call’ and ‘Occupied’ statuses

(lighter colour represents expected position in the next decision epoch)

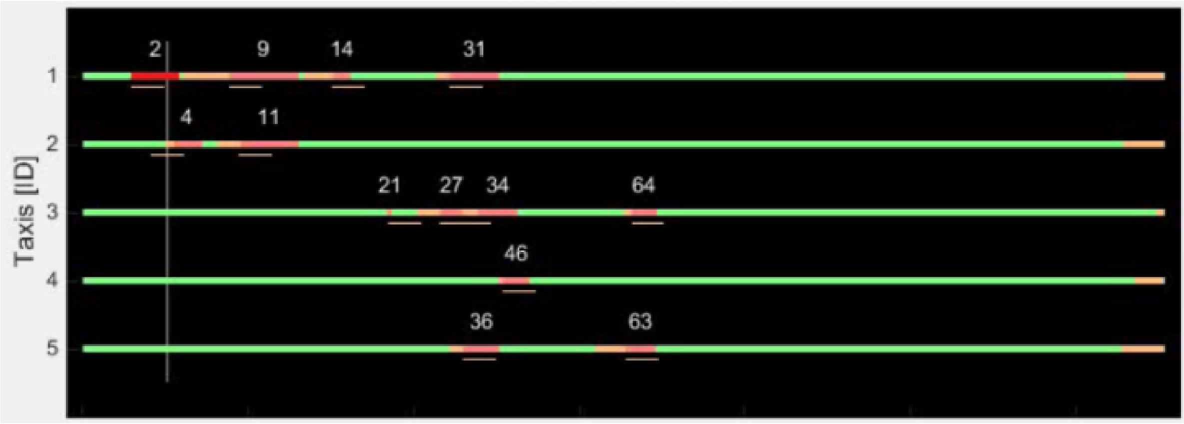
Passenger origin and destination location specified in a booking request (in orange is a passenger waiting for a taxi)

Visualization of taxi assignment to pick up particular passenger

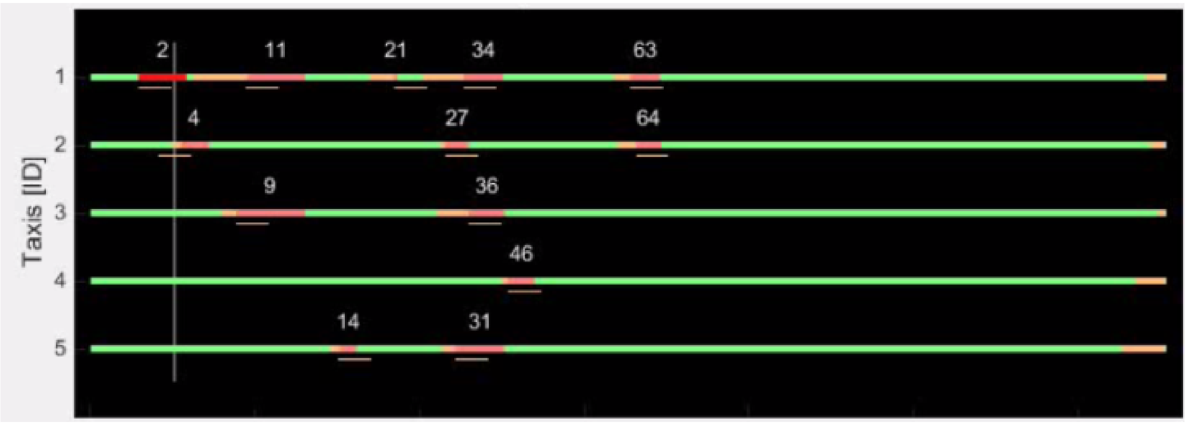
Figure 33 Taxi simulation screenshot – spatial dimension

Temporal dimension – taxi schedules

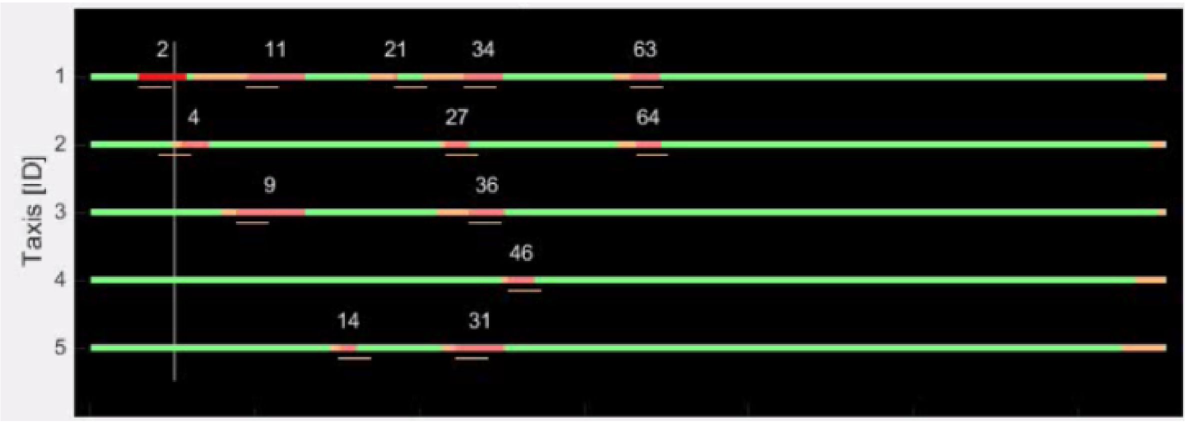
Schedule after insertion of new booking requests



Schedule after re-assignment of confirmed booking requests



Accepted schedule



Current time → Time

Legend

- 'Vacant' time
- 'Occupied' time
- 'On-call' time

Figure 34 Taxi simulation screenshot – temporal dimension

Table 17 summarizes the statuses of taxis and booking requests. From the perspective of the simulation, the statuses have impact on the future, current runtime or for gathering the statistics about the past events.

	Future	Current runtime	Past statistics
Booking request	'Initial'	'Booked' 'Confirmed' 'Waiting' 'On-trip'	'Booking failed' 'Waiting failed' 'Delivered'
Taxis	'Off'	'Vacant' 'On-call' 'Occupied'	

Table 17 Statuses of taxi and passenger requests

Figure 35 shows an example of the communication dispatching protocol of the simulation. In this example, the simulation executes two new booking requests. These requests are assigned to the best suiting taxis and confirmed to the passengers. In the following re-assignment phase, groups of booking requests are re-assigned (in this case coincidentally the same ones). The total costs and total waiting time of the schedule before and after the re-assignment are compared and the more favourable accepted. Upon this decision, the schedule is updated. The changes may have an impact on other requests in the schedule, so that any requests, which are impacted by the re-assignment, must be updated in terms of estimated pick-up times, drop-off and other variables. Finally, the simulation schedules all known events in the future.

Time	Explanation	Dispatching communication log
06:02:15	New booking requests (5102 14770) are assigned	NEW REQESTS: 5102 and 14770 requested booking. ASSIGN REQUEST 5102 Estimated 'on-call': 06:02:20 pick-up: 06:02:27 drop-off: 06:27:56 by taxi T 174. ASSIGN REQUEST 14770 Estimated 'on-call': 06:02:20 pick-up: 06:02:26 drop-off: 06:26:44 by taxi T 92.
	Re-assignment	RE-ASSIGNMENT: A batch of confirmed trips: (14770, 5102)
	Comparing costs and waiting time	Costs_Existing :-0.014069 $ WaitingTime_Existing: 31 s Costs_SM: -0.009626 $, WaitingTime_SM: 27 s
	Acceptance	RE-ASSIGNMENT is accepted for this batch.
	Taxi schedule housekeeping	Trip 14770 removed from the schedule of taxi T 92 Trip 5102 removed from the schedule of taxi T 174 Trip 5102 will be picked up by taxi: T 404 at: 06:02:27 Trip 14770 will be picked up by taxi: T 174 at: 06:02:22
06:02:20	Departures, pick-ups and new passenger waiting housekeeping	DEPARTURE: Taxi T 174 will depart to trip 14770 at: 06:02:20 DEPARTURE: Taxi T 404 will depart to trip 5102 at: 06:02:20 PICK-UP: Taxi T 174 will pick-up trip 14770 at 06:02:22 WAITING: Passenger 14770 will begin waiting at 06:02:20 WAITING: Passenger 5102 will begin waiting at: 06:02:20

Figure 35 An example of the communication dispatching protocol

Figure 36 shows an example of an event log of a taxi. Taxis begin a shift by switching the status from 'off' to 'vacant'. If they are on a way to pick up a confirmed passenger, they have status 'on-call', which changes to 'occupied' as the passenger boards the taxi and to 'vacant' after reaching passengers destination.

Time	Taxi event log
06:01:54	VACANT T 174 has begun a shift.
06:02:20	ON-CALL T 174 is 'on-call' and departed to passenger 14770.
06:02:22	OCCUPIED T 174 is 'occupied' with passenger 14770.
06:26:40	VACANT T 174 dropped off passenger 14770

Figure 36 An example of a taxi event log

Figure 37 shows the simplest process for the passenger. First passenger requests a taxi, and then gets the confirmation. After the passengers starts waiting and eventually the taxi comes to pick up the passenger and drivers to the destination drop-off.

Time	Passenger request event log
06:02:11	REQUEST 5102 received. We will reply in max 0.16667 min from now.
06:02:15	UPDATE 5102 'on-call': 06:02:20, pick-up: 06:02:27, drop-off: 06:27:56 by T 174.
06:02:20	WAITING 5102 You have just started waiting for the assigned taxi.
06:02:27	PICK-UP 5102 You were just picked up by taxi T 404.
06:27:56	DROP-OFF 5102 You have reached your destination.

Figure 37 An example of a passenger (trip) event log

Figure 38 shows an example of the greediness aspect and, in particular, displays a passenger, who experiences two re-assignments. First, the estimated pick-up time reduces by about four minutes. Second, because a taxi managed to add one other passenger in between, the estimated pick-up time is shifted by about one minute. This example is important to understand, the greediness of the passengers (and similarly taxis) in accepting worse assignments than they are actually currently having.

Time	Passenger request event log
14:29:05	REQUEST 14120 received. We will reply in max 0.16667 min from now.
	UPDATE 14120 'on-call': 14:43:30, pick-up: 14:45:24, drop-off: 14:46:16 by T 182.
14:29:10	WAITING 14120 You have just started waiting for the assigned taxi.
14:30:20	UPDATE 14120 Caused by this trips insertion. 'on-call': 14:39:19 pick-up: 14:41:13 drop-off: 14:42:05 by T 182.
14:39:45	UPDATE 14120 changes of (12920) whose 'on-call': 14:39:50, drop-off: 14:42:00 AFFECTS 14120, so that 'on-call': 14:42:00 pick-up: 14:42:52, drop-off: 14:43:44 by T 182.
14:42:52	PICK-UP 14120 was just picked up by taxi T 182.
14:43:44	DROP-OFF 14120 has reached your destination.

Figure 38 A passenger event log with improvement and worsening of the estimated pick-up time.

Table 18 shows an example of five taxis and their variables at some time of the simulation. Some taxis are not in service, and those, who are, are either ‘vacant’, or traveling to the passenger or serving the passenger.

ID	Status	Depot position		Shift	
		Lat.	Lon.	begin	end
T 1	'Vacant'	-122.41486	37.78094	1212386428	1212407758
				1212419240	1212450060
				1212463327	1212472793
T 2	'Off'	-122.40527	37.79812	1212436545	1212463870
				1212468943	1212472767
T 3	'On-call'	-122.4108	37.76817	1212386423	1212400969
				1212403088	1212469878
T 4	'Occupied'	-122.41882	37.80752	1212386447	1212395003
				1212413524	1212420393
				1212424776	1212426372
				1212429560	1212465593
				1212471016	1212472748
T 5	'Occupied'	-122.41257	37.80828	1212386450	1212403242
				1212414943	1212472772

Table 18 Example of taxi variables (1/2)

Table 19 shows detailed history of taxi ‘on-call’ movements as well as history of served passenger trips. In both cases it shows the ID of the booking request and the actual begin and end of the status. One can also see, the link with the previous figure. For example taxi number 3 has status ‘on-call’ and that’s why the variable ‘on-call’ time end as well as the on-trip variables are yet empty. Similarly taxis 4 and 5 are still on the way with passengers at the moment and therefore the variable on-trip time end in the table is also empty.

ID	‘On-call’			‘Occupied’		
	Trip ID	begin	end	Trip ID	begin	end
T 1	[10715, 7262, 13607, 15523]	[1212389000, 1212389880, 1212390405, 1212391845]	[1212389017, 1212389887, 1212390416, 1212391907]	[10715, 7262, 13607, 15523]	[1212389017, 1212389887, 1212390416, 1212391907]	[1212389185, 1212389990, 1212390575, 1212392135]
T 2	[8381, 15054, 2512]	[1212386905, 1212390925, 1212391422]	[1212386931, 1212390926, 1212392288]	[8381, 15054, 2512]	[1212386931, 1212390926, 1212392288]	[1212387118, 1212391422, 1212393721]
T 3	[6529, 6330, 41, 14816, 15979]	[1212387220, 1212392905, 1212393185, 1212393415, 1212393848]	[1212387246, 1212392616, 1212392919, 1212393202, 1212393444]	[6529, 2289, 6330, 41, 14816]	[1212387246, 1212392616, 1212392919, 1212393202, 1212393444]	[1212387355, 1212392775, 1212393115, 1212393286, 1212393848]
T 4	[3578, 3037]	[1212386450, 1212392355]	[1212386463, 1212393460]	[3578, 3037]	[1212386463, 1212393460]	[1212386592]
T 5	[10903, 7083, 4751, 1360, 6757]	[1212386545, 1212390300, 1212392669, 1212392895, 1212393607]	[1212386559, 1212390344, 1212392686, 1212392934, 1212394563]	[10903, 7083, 4751, 1360, 6757]	[1212386559, 1212390344, 1212392686, 1212392934, 1212394563]	[1212386683, 1212390449, 1212392839, 1212393607]

Table 19 Example of taxi variables (2/2)

Similarly, Table 20 shows the key variables of passengers: The ID, status, desired pick-up time interval, type of booking, time of booking and latitude and longitude of origin and destination.

ID	Desired pick-up begin	Desired pick-up end	Booking time	Origin position Lat.	Origin position Lon.	Destination position Lat.	Destination position Lon.
T 1	1212437729	1212439529	1212437729	-122.42036	37.78119	-122.40743	37.79206
T 2	1212393751	1212395551	1212393751	-122.39092	37.61546	-122.46297	37.73179
T 3	1212386822	1212388622	1212386822	-122.43375	37.77175	-122.47396	37.76370

Table 20 Example of passenger variables (1/2)

Table 21 than shows other variables of the passengers, which show the expected and actual events, most importantly the pick-up and drop-off times. Also here a connection between Table 20 can be seen. Because request 1 is yet in the initial status (has not booked yet), the variables are yet empty. Request 2 is 'On-trip' so that's why the actual drop-off time is still missing. So far everything went as planned. Trip 3 is an example of a deliver trip.

ID	Status	Trip ID	Estimated time of On-call	Estimated time of Pick-up	Estimated time of Drop-off	Actual time of Pick-up	Actual time of Drop-off	Waiting
T 1	'Initial'	[]	[]	[]	[]	[]	[]	[]
T 2	'On-trip'	166	1212393814	1212394844	1212395940	1212394844	[]	1093
T 3	'Delivered'	490	1212386830	1212386837	1212387115	1212386837	1212387115	15

Table 21 Example of passenger variables (2/2)

3.5 Performance indicators analysis

As the review of performance indicators revealed, various authors use various indicators and there is no generally accepted consensus (unlike the classical VRP benchmarks) as of which indicators should be used. Therefore, this study aims to sort this situation partly by relating those indicators that have actually some dependencies.

Tables 9 – 11 in the literature review section 2.4 about experimental design reveal the following: (1) As far as passengers are concerned, the waiting time until a taxi comes to pick up is the most mentioned indicator. (2) For taxis, one of the dominant indicators is the number of taxis required to satisfy the given demand, but because in the experiments fixed fleet is assumed, this indicator is not relevant in that sense. As next, costs and profits are mentioned very frequently. Furthermore, time or distance based variables such as 'vacant' and 'occupied' distance or time are also quite frequent. Some other less frequent rates give an indication that the ratio between occupied and total time or distance is for some a measure of system efficiency and utilization. (3) From the dispatching system point of view, the most frequent and tangible at the same time is the number of the satisfied booking requests. Taking the above mentioned in consideration and in order to describe the dispatching system performance, this study selected the following performance indicators:

- Average passenger waiting time – the only indicator with wider consensus among the research studies and the only indicator describing the passenger expectations and inconvenience. The passenger waiting time begins at the moment the passenger begins booking a taxi and ends when taxi picks up the passenger.
- The number of served passenger booking requests – indicator that describes not just the ability of the dispatching system to process booking requests, but also the passenger satisfaction. Most of the dispatching systems may reject a passenger request, if there is high passenger demand and not sufficient taxi supply.
- Average 'on-call' taxi distance and average total distance – these two indicators provide a physical dimension to the economic description of taxi profit as well as give an indication about taxi utilization. The 'on-call' distance is the distance taxi travels when assigned to pick up a particular passenger.
- Average taxi profit – this indicators can be seen as a tip of an iceberg, because it integrates many other indicators. The dependencies of the other variables are described in the paragraphs below.

Methodology of enumeration of costs and profit

Costs and profits associated with picking up a passenger requests are abstracted. In detail, Figure 39 shows the situation of a 'vacant' taxi searching for a passenger and evaluating the various utilities from a time perspective. Table 22 accompanies Figure 39 with the variables in full together with their units.

During the pre-trip time (before pick-up) costs only incur to the taxi driver. There are three main sources of costs: (1) the costs of driving a distance, (2) the costs of the driver's time and vehicle rental or purchase, and (3) the costs triggered by events (for example tolls). All fixed costs (such as insurance, regular maintenance, road tax) are prorated by the company through the vehicle rental or purchase costs.

There are also three main sources of revenues: (1) revenues from distance fares, (2) revenues from waiting time, and (3) revenues triggered by events (for example location surcharges or any applicable ancillary charges). The taxi driver's profit is what remains from the revenue after subtracting costs over an observed time of one shift. The profit earning rate expresses the rate of profit earing, or in other words, how fast the driver earns in time. (Kümmel et al., 2015)

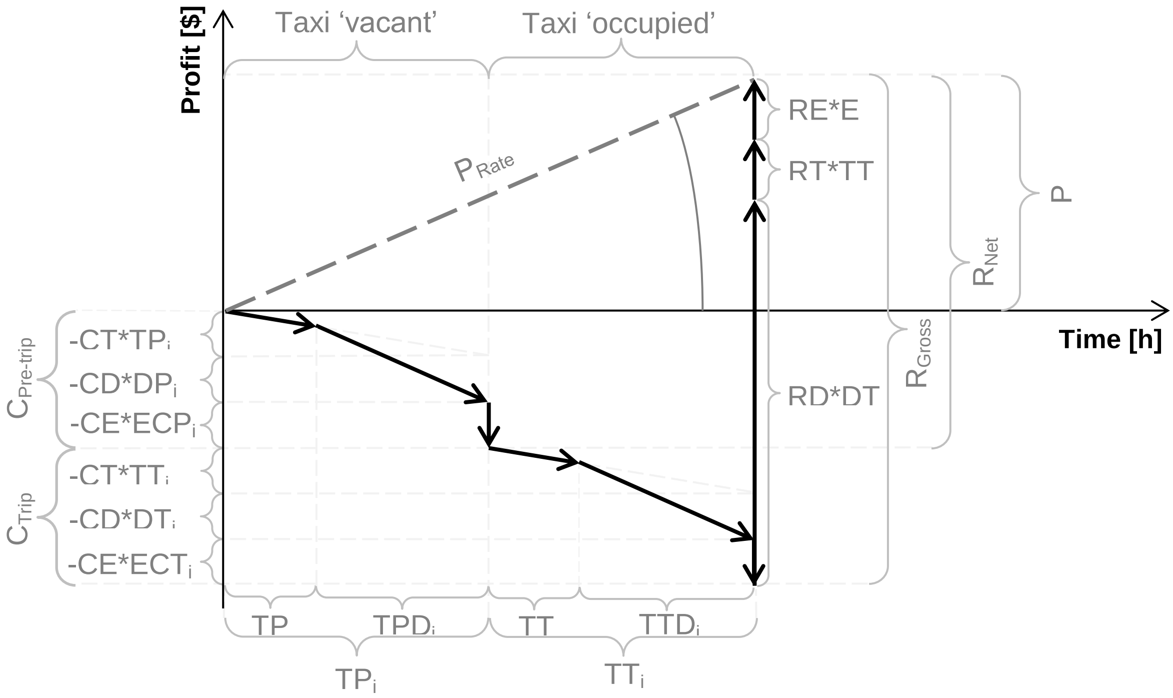


Figure 39 Calculation of booking request's costs and revenue (Kümmel et al., 2015)

Ab.	Parameter	Unit	Ab.	Parameter	Unit
P	Profit	[$]	DPi	Distance to pick-up location	[km]
P_{Rate}	Profit earning rate	[$/h]	DT_i	Trip distance (from O_i to D_i)	[km]
R_{Gross}	Gross revenue	[$]	TP_i	Pre-trip time	[h]
R_{Net}	Net revenue	[$]	TPS_i	Pre-trip time standing	[h]
RD	Revenue rate from distance	[$/km]	TPD_i	Pre-trip time driving	[h]
RT	Revenue rate from time	[$/h]	TT_i	Trip time	[h]
RE	Revenue rate from event	[$/event]	TTS_i	Trip time standing	[h]
C_{Trip}	Trip costs	[$]	TTD_i	Trip time driving	[h]
$C_{Pre\text{-}Trip}$	Pre-trip costs	[$]	ER_i	Revenue event count	[1]
CD	Cost rate of distance	[$/km]	ECP_i	Cost event count pre-trip	[1]
CT	Cost rate of time	[$/h]	ECT_i	Cost event count trip	[1]
CE	Cost rate of event	[$/event]	TV	Traveling velocity	[km/h]

Table 22 Parameters, their abbreviations and units (Kümmel et al., 2015)

The profit (P) consists of revenues (R) and costs (C). Both revenues and costs can be triggered by either distance, time or an event. The average taxi profit indicator is defined as follows.

$$\overline{P} = \sum_{j=1}^{n(Taxis)} P_j \Big/ n(Taxis) \quad (1)$$

$$P_j = R_j - C_j \quad (2)$$

The revenues are determined by the fares per distance (RD) and the sum of the occupied trip distance driven (DST), by fare per time (RT) and the sum of the occupied standing time (TST), in which passenger is on board but the vehicle is not moving due to other traffic or traffic lights and by fare surcharge events such as booking fees or location fees (RE) and their number (ER), all calculated over the set of occupied trips (SOTj).

$$R_j = RD * \sum_{i \in SOT_j} DST_i \; + \; RT * \sum_{i \in SOT_j} TST_i \; + \; RE * \sum_{i \in SOT_j} ER_i \tag{3}$$

The costs are determined by the cost per distance (CD) and the sum of all distance driven (DT) on all 'occupied' and 'on-call' trips, by the cost per time (CT) and sum of all hours when the vehicle is active (TT) and by the cost of events such as city centre congestion charges and tolls (CE) and their occurrence (EC), all calculated over the set of all trips (SATj).

$$C_j = CD * \sum_{i \in SAT_j} DT_i \; + \; CT * \sum_{i \in SAT_j} TT_i \; + \; CE * \sum_{i \in SAT_j} ECT_i \tag{4}$$

The number of served passenger booking requests (pick-ups) reflects the capability of the given fleet and matching algorithm to deliver the service. The average 'on-call' distance is a sum of the distances driven to the passengers (DCT) when the status of the taxi is 'on-call'. This indicator shows the influence of group simultaneous assignment in reducing the 'on-call' distance and related costs over the set of 'on-call' trips (SOCTj). The average total distance is a sum of the distances driven by the taxis (DT) when the taxi is in service (i.e. when the status is anything other than 'Off'). Passenger waiting time indicates the time elapsed from the desired pick-up time (PT) to the actual pick-up time (APT). If the passenger is not picked up before the expiration of the passenger's acceptable waiting period (WW), the waiting time equals the acceptable waiting period. (Kümmel et al., 2016c)

$$n(PickUps) = \sum_{j=1}^{n(Taxis)} n(SOT_j) \tag{5}$$

$$\overline{D}_{On-call} = \sum_{j=1}^{n(Taxis)} \sum_{i \in SOCT_j} DCT_i \Big/ n(Taxis) \tag{6}$$

$$\overline{D}_{Total} = \sum_{j=1}^{n(Taxis)} \sum_{i \in SAT_j} DT_i \Big/ n(Taxis) \tag{7}$$

$$\overline{T}_{CustomersWaiting} = \sum_{i=1}^{n(Customers)} \begin{cases} APT_i - PT_i, & APT_i \neq \{\} \\ WW_i, & APT_i = \{\} \end{cases} \Big/ n(Customers) \tag{8}$$

4 Simulation experiments and results

In order to find the answers to the three research questions and to analyse the impacts of the proposed dispatching methodological approaches three experiments were made. Tables 23 to 27 extensively summarize all parameters used in the three experiments and their values used for simulation, covering the variables related to demand, supply, dispatching algorithm, environment and the simulation model in each of the three experiments. This chapter describes the experiments and the respective results in detail.

Demand parameters	**Experiment 1** **Simultaneous assignment of taxis to new immediate bookings**	**Experiment 2** **Re-assignment of all confirmed immediate and advance bookings**	**Experiment 3** **Re-assignment of the timely nearest confirmed immediate bookings**
Time interval, during which passenger requests can come in [h]	4	8	24 (Monday 2008-06-02 06:00:00 – Tuesday 2008-06-03 06:00:00)
Passenger origin and destination positions	Uniformly distributed	Uniformly distributed	Based on real data shown in Figure 53
Passenger desired pick-up times	Uniformly distributed	Distribution in Figure 45	Based on real data shown in Figure 55
Passenger booking times	Same as desired pick-up time	Immediate bookings: Same as desired pick-up time Advance bookings: Uniformly distributed between the simulation beginning and desired pick-up time – 15 min	Same as desired pick-up time
The number of booking requests	1200	2400, 1600 and 800	16302
Average number of booking requests per hour	300	300, 200 and 100	679
Average booking requests per hour per taxi	3	3, 2 and 1	1.3
Passengers' willingness to wait [s]	1000	1000	1000
Degree of dynamism	100% Just immediate bookings	<0, 10, 20,.. ,100% >	100% Just immediate bookings
Advance booking time	Not applicable	15 min and more ahead	Not applicable
Status of bookings at the beginning of simulation	'Initial'	'Initial'	'Initial' or 'booked'

Table 23 Comparison of experiment setups – demand parameters

Supply parameters	**Experiment 1**	**Experiment 2**	**Experiment 3**
Taxi positions	Uniformly distributed (See Figure 40)	Uniformly distributed See Figure 40)	Based on real data
Taxi availability	From the beginning until the end of simulation	From the beginning until the end of simulation	In predefined shifts according to real shifts
The number of taxis	100	100	536
Status of taxis at the beginning of simulation	'Vacant'	'Vacant'	'Vacant' or 'off'

Table 24 Comparison of experiment setups – dispatching system parameters

Dispatching parameters	**Experiment 1**	**Experiment 2**	**Experiment 3**
Buffering time interval length [s]	30	60	5
Length of re-assignment interval	Not applicable	Until the simulation end	1, 2, 3.75, 5, 7.5, 15, 20, 30 and 60 min
Which taxi participate the initial insertion?	'Vacant'	Any	Any
First assignment	Simultaneous	Sequential	Sequential
Order of the new bookings	Not applicable	Bookings sorted according to booking time	Bookings sorted according to booking time
Taxi preferences for assignment	Taxi costs (time and distance)	Not applicable	Not applicable
Passenger preferences for assignment	Estimated pick-up time	Not applicable	Not applicable
Re-assignment method	Not applicable	Simultaneous	Simultaneous
Taxi preferences for re-assignment	Not applicable	Additional costs of inserting to schedule	Change of 'on-call' distance costs
Passenger preferences for re-assignment	Not applicable	Estimated pick-up time	Estimated pick-up time. If trips or taxis are greedy, all preferences worse than the currently engaged trip or taxi are not included.
Which party proposes?	Taxis	Taxis and passengers	Passengers
Order of proposals	Random	If taxis propose, in each round, first the taxi, which has not proposed yet and has the highest additional costs to pick up a trip. If passengers propose, the passenger, who has not yet proposed and has the lowest estimated pick-up time.	First proposers selected according to the least number of possible matches, according to the desired pick-up time begin and then randomly.
Prediction of taxis at the next interval	No	Linear estimation of taxi positions and statuses	No
Prediction length	Not applicable	Decision epoch length	Not applicable
Simulated assignment duration [s]	15	Not considered	Not considered
If booking rejected	Repeat booking until the expiry of initial willingness to wait	Repeat booking until the expiry of initial willingness to wait	Do not try again

Table 25 Comparison of experiment setups – supply parameters

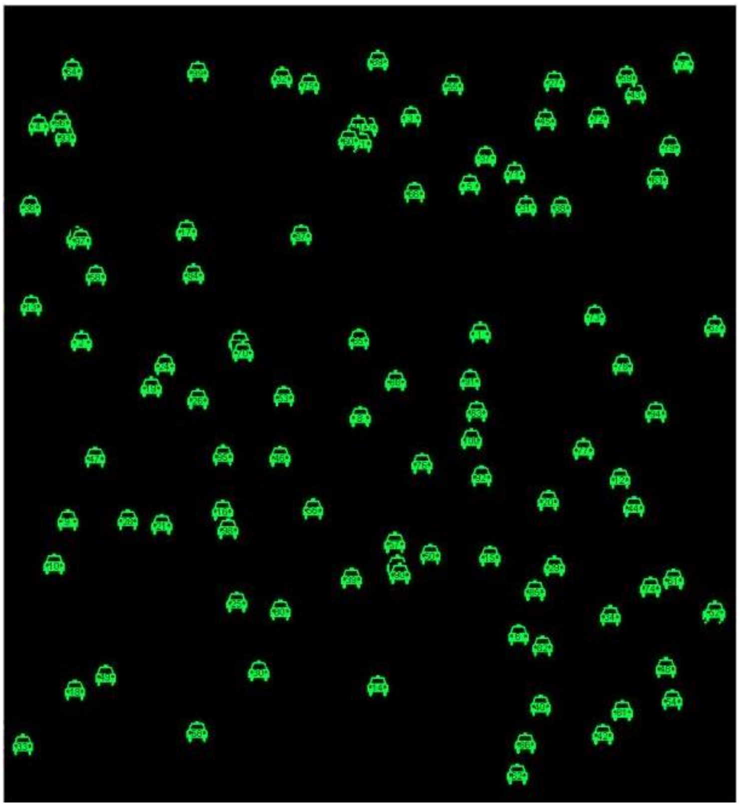


Figure 40 Example of initial positions of taxis in Experiment 1 and 2

Environment parameters	Experiment 1	Experiment 2	Experiment 3
Size of area	Square of size 20 km	Square of size 20 km	Lat. <37.4, 38.1>, Lon. <-122.6,-121.9 >
Velocity of taxis [km/h]	36	36	13.158
Waiting strategy	Move → wait	Wait → move	Wait → move

Table 26 Comparison of experiment setups – environment parameters

Simulation parameters	Experiment 1	Experiment 2	Experiment 3
Simulation ends	After there are no more events to be executed in the simulation plan	At pre-set time (after even the latest trip can be served)	Tuesday 2008-06-03 06:00:00 in Unix time)

Table 27 Comparison of experiment setups – simulation model parameters

Table 28 summarizes the assumed variable values common to all experiments. These values are a result of an educated guess, stemming from various sources, with the objective to get a reasonable estimation of the costs, revenues and other variables.

Abbreviation	Parameter	Value	Unit
RD, CD	Revenue and cost rate per distance	1.560, 0.071	[$/km]
RT, CT	Revenue and cost rate per time	19.560, 8.333	[$/h]
CE, RE	Revenue per event, cost per event	Not considered	[$/event]

Table 28 Input assumptions common to all experiments

4.1 Simultaneous assignment of taxis to new immediate bookings

The objective of the first experiment is to analyse the effects of replacing the conventional first-come, first-served dispatching strategy by simultaneous assignment of new booking requests, as discussed in the methodological part 3.1.3. “Simultaneous assignment of taxis to immediate booking requests”. Moreover, this experiment introduces three factors that extend the existing simple study by Bai et al. (2014) in the quest for more accurate assumptions and more efficient dispatching algorithms.

The introduced factors create new variants of the dispatching strategy. Stable marriage strategy variants with subsequent adding the three factors of limited passengers’ willingness to wait, commitment to already confirmed requests and appropriate length of decision epoch are further referred as: SM_0, SM_1, SM_2 and SM_3 and summarized in Table **29**. First-come, first-served strategy serves as a benchmark.

Factor	**FCFS**	**SM_0**	**SM_1**	**SM_2**	**SM_3**
Limited willingness to wait	✓	✘	✓	✓	✓
Commitment to confirmed trips	✓	✘	✘	✓	✓
Appropriate length of decision epoch	n.a.	✘	✘	✘	✓

Table 29 Stable marriage based strategies with extension factors (Kümmel et al., 2016d)

As summarized in Table 23 and Table 24, 100 taxis and 1200 passenger immediate booking requests generated during 4 hours are simulated. That corresponds to high demand of an average of 3 booking request per taxi and per hour. The booking times are uniformly distributed. The decision epoch length is set to 30 s. The assignment method is taxi centric. In order to minimize the influence of the initial set-up on the results, all strategies are simulated in 15 simulation rounds. These 15 spatial-temporal configurations determine the initial taxi positions, origins and destination of the trips and request times. The presented results are based on the mean of these experiments.

The results are summarized in Tables 30 and 31, Figure 41 and commented bellow. In the custom-made simulation, the FCFS strategy serves on average 1064 passengers (rounded to integers) out of 1200 (88.7%) in less than 13 minutes of average waiting time. Herewith, in the perspective of rather high trip demand (3 passenger requests per hour per taxi), the FCFS strategy sets a high benchmark.

SM_0 does not limit willingness to wait, which results in both advantages and disadvantages. On one hand, this strategy serves all 1200 passenger trips (100%) and produces the highest taxi profit among all compared strategies (see Tables 30 and 31). On the other hand, passengers have to wait almost 20 minutes on average to get a taxi, which might not be acceptable by many passengers.

SM_1 strategy limits willingness to wait. It manages to serve 1027 passengers (85.6%). The average waiting time and total distance are the best among all the compared strategies. However, this strategy poses a disadvantage; periodically generated assignment of taxis and passengers might be inconsistent between decision epochs. In strategy SM_1, unlike in the others, the number of not picked up but confirmed passengers reached 26, corresponding to 2.5% of passengers.

SM_2 strategy forbids changing of confirmed assignment and indeed all passengers with confirmed booking are picked up. However, this is compensated with lower taxi driver profit and increased not-occupied and total distance. This strategy serves 1033 trips (86.1%). In direct comparison with FCFS strategy, it improves not-occupied and total distance but worsens the number of served passengers and profit. This means that, while helping the passengers, city traffic and the environment, this strategy does not really favour taxi drivers and operators. Hardly any manager or board member in taxi operating company would accept a profit-reducing strategy. (Kümmel et al., 2016b)

Performance indicator	FCFS	SM_0	SM_1	SM_2	SM_3
Average taxi profit [$]	93.5	107.9	90.4	90.5	97.6
No. of served bookings [1]	1063.8	1200.0	1027.4	1033.2	1089.4
Average 'vacant' taxi distance [km]	37.5	20.5	21.1	22.3	30.5
Average total taxi distance [km]	147.5	145.5	127.5	129.0	143.2
Average passenger waiting time [mm:ss]	12:40	20:00	08:27	08:53	08:36

Table 30 Mean of performance indicators (Kümmel et al., 2016d)

Performance indicator	FCFS	SM_0	SM_1	SM_2	SM_3
Average taxi profit [$]	1.5	1.8	1.4	1.6	1.4
No. of served bookings [1]	9.8	0.0	11.9	11.6	10.6
Average 'vacant' taxi distance [km]	1.2	1.7	0.6	0.7	1.3
Average total taxi distance [km]	1.7	1.8	1.3	1.5	1.5
Average passenger waiting time [mm:ss]	00:10	01:15	00:08	00:07	00:12

Table 31 Standard deviation of performance indicators (Kümmel et al., 2016b)

Performance indicator		SM_0	SM_1	SM_2	SM_3
Average taxi profit		15.4%	-3.4%	-3.3%	4.4%
No. of served bookings		12.8%	-3.4%	-2.9%	2.4%
Average 'vacant' taxi distance		45.4%	43.8%	40.6%	18.6%
Average total taxi distance		1.3%	13.5%	12.5%	2.9%
Average passenger waiting time		-57.9%	33.3%	29.8%	32.2%

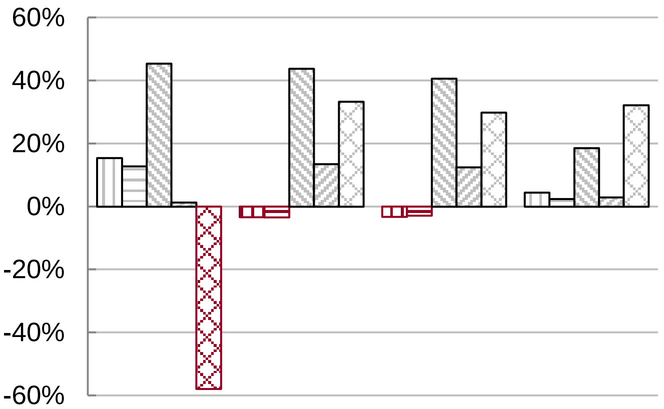


Figure 41 Positive values depict improvement over FCFS (Kümmel et al., 2016b)

SM_3 strategy, thanks to appropriate length of decision epoch, serves 1089 trips (90.1%). The number of served trips increases by 2.4%, profit increases by 4.4%, the highest among the strategies with limited willingness to wait. All other observed parameters also improved as compared to FCFS strategy; not-occupied and total taxi distance and passengers' waiting time by 18.6%, 2.9% and 32.2% respectively. Hence, taxis serve more passengers and collect higher profits whilst reducing empty and total driven distance. Moreover, the passengers benefit from waiting time reduction. Finally, thanks to reduced distance driven, the urban environment receives less traffic and associated pollution. (Kümmel et al., 2016b)

Figure 42 provides a detailed view on the results of Experiment 1. In particular, it shows the results of all five dispatching strategies in all five performance indicators on the level of individual taxis and passengers. The results of each taxi or passenger in each category and each indicator are ordered from the highest to the lowest and an average of the values is highlighted by a dashed line. Apart from the distribution of the results across individual taxis, there are couple of other observations to make. For example, the results claim the strategy SM_0 picks up all passengers. Observe, what it means for their waiting time (last row, second figure from the left). Because these figures are cropped at the same value of about 33 min of waiting time, some results are not shown since they are higher than that. The strategies SM_1 to SM_3 have cap on willingness to wait (16.7 min), which also reflects in the actual waiting time.

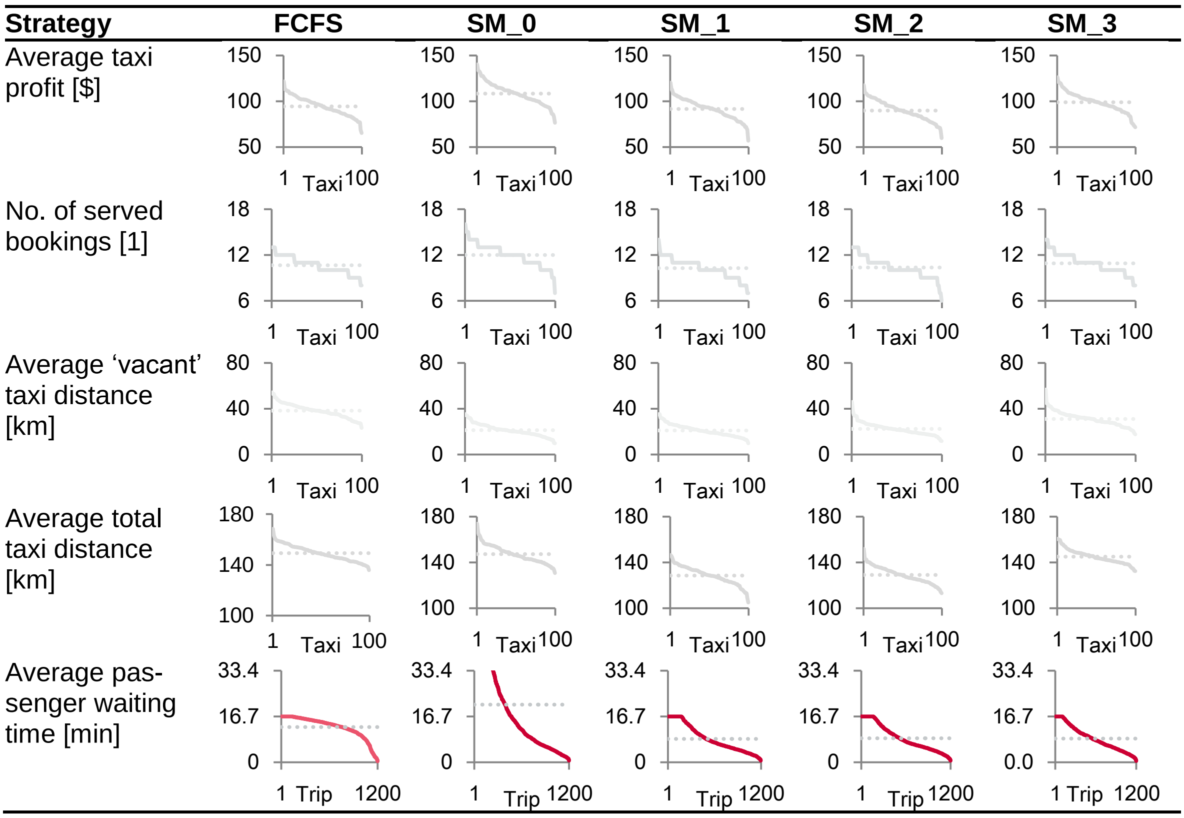


Figure 42 Detailed view of the results of Experiment 1

A sensitivity analysis of the simulation environment is carried out based on the FCFS strategy. Following input parameters are varied by +-10%: Travel velocity, willingness to wait, length of a square area side, the number of taxis and customers. The relative increase caused by varying input parameters is summarized in Table 32.

The sensitivity analysis showed that only variation of moving velocity and willingness to wait causes relative change higher than 10%. Increased velocity increases average 'vacant' distance by 16%. Decreased velocity decreases 'vacant' distance by 19%. These observations are in line with analysis of taxi GPS logs by Liu et al. (2010), which concluded that higher moving velocities indeed lead to a higher income for taxi drivers. Increased willingness to wait increases average waiting time by 12%. Reduction of willingness to wait reduces waiting time by 13%.

To sum up, the sensitivity analysis of the simulation environment showed consistent results in line with common sense expectations. Thereby the model can be affirmed as suitable for evaluation of impacts of dispatching strategies.

Parameter	**TV**		**WW**		**SS**		**n(Taxis)**		**n(Customers)**	
Variation (+-%)	**+10**	**-10**	**+10**	**-10**	**+10**	**-10**	**+10**	**-10**	**+10**	**-10**
Average taxi profit	10%	-9%	0%	0%	6%	-6%	-4%	5%	5%	-6%
No. of served bookings	6%	-7%	1%	0%	-5%	6%	7%	-8%	3%	-4%
Average 'on-call' taxi distance	16%	-19%	2%	-3%	-8%	9%	3%	-6%	-6%	9%
Average total taxi distance	9%	-10%	1%	-1%	1%	-1%	-1%	0%	1%	-1%
Average passenger waiting time	-4%	3%	12%	-13%	3%	-4%	-5%	4%	5%	-8%

Table 32 Sensitivity analysis FCFS strategy

The trade-off between the length of batching interval and the system responsiveness is investigated thoroughly. In particular, Figure 43 shows the results of varying the length of decision epoch (the buffering interval).

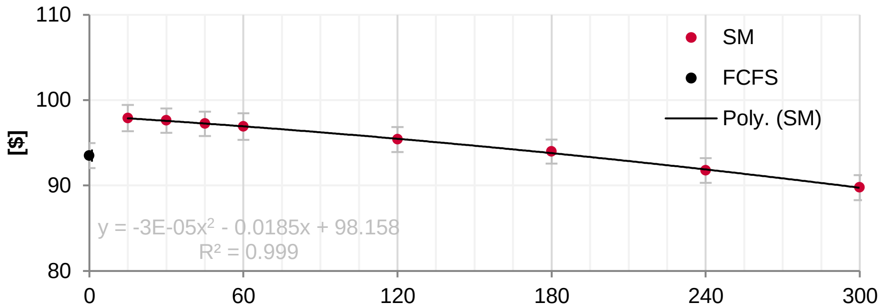
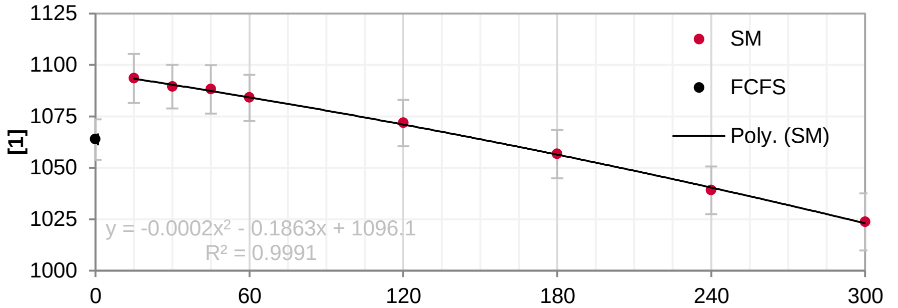
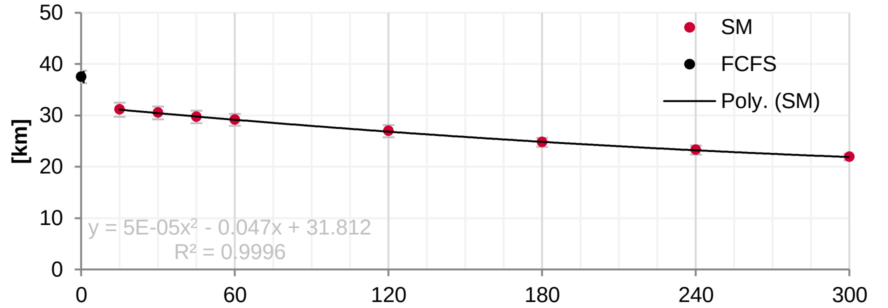
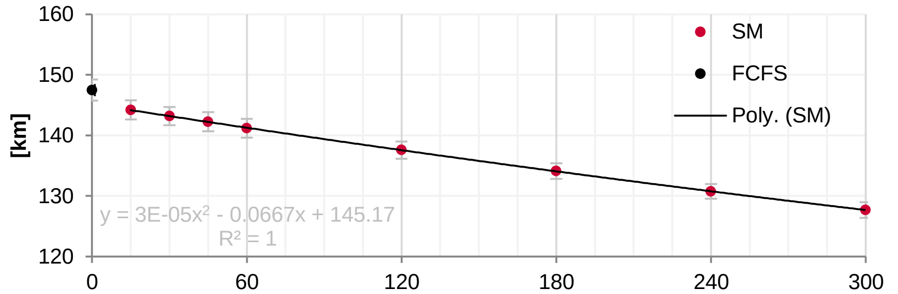
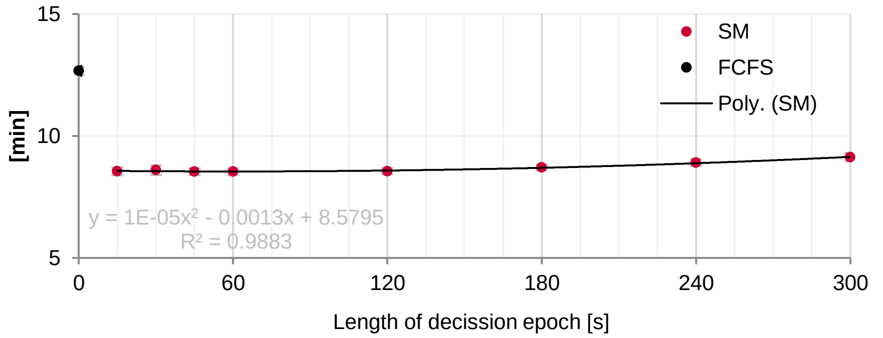
<table>
<tr><th>Indicator</th><th>Sensitivity analysis of varying decision epoch of SM_3 on the results</th></tr>
<tr><td>Average taxi profit</td><td>
</td></tr>
<tr><td>No. of served bookings</td><td>
</td></tr>
<tr><td>Average ‘on-call’ taxi distance</td><td>
</td></tr>
<tr><td>Average total taxi distance</td><td>
</td></tr>
<tr><td>Average passenger waiting time</td><td>
</td></tr>
</table>

Figure 43 Sensitivity analysis of varying decision epoch of SM_3 on the results

Summary of methodology of dispatching taxis to immediate booking requests

To sum up, this chapter summarizes the application of the stable marriage algorithm for assigning immediate taxi booking requests to taxis. Double-sided simultaneous assignment strategy, which caters for preferences of both taxis and passengers, is able to cope with the unequal number of pairs and preference ties. Main limitation of this approach is that only immediate booking requests not advance booking requests are considered. Well it is not impossible to assign them simultaneously, but Approach 1 lacks any re-assignment mechanism that would prevent the worsening of taxi schedules after new bookings will come in.

Thy hypothesis: “The potential of simultaneous assignment to improve taxi dispatching performance increases with the amount of booking requests that can be assigned at the same time and works the best for immediate bookings.” can be confirmed. As mentioned above, it is not suitable for advance bookings, because of the lack or re-assignment. As mentioned earlier, the more the booking requests at similar time and at similar position, the higher are the chances that re-assignment procedure will find more suitable assignment of taxi to passengers.

Estimating the costs of stability

In order to get a sense for the price of stability of the stable marriage algorithm and how the solution performs as compared to the optimal solution, it is worthwhile to compare it with a ground truth by using the brute force approach. The brute force approach means that all possible combinations of assignment of taxis to passengers are evaluated in terms of costs and / or waiting time. A complexity of O(n!) means that, for instance, for 2 passengers and 2 taxis there are 2 possible assignment combinations, for 3 passengers and 3 taxis 6 combinations but for 10 passengers and 10 taxis 3.6x10^106 and for 100 passengers and 100 taxis 9.3x10^10157. And this is yet without any pre-scheduled passengers or trip sequences, which would increase the complexity even further.

In order to illustrate at least roughly and in small scale the performance of the stable matching as compared to all brute force solutions, the following short experiment is made. The stable marriage algorithm implementation as in Experiment 1 is used to generate assignments of the same number of taxis and passengers. Figure 44 summarizes all solutions ordered according to the total costs of pick-up a passenger represented as black dots. The red dot represents the results of the stable assignment. The closer is the dot to the first black dot on the optimal solution on the very left; the smaller is the gap and the price of stability. This experiment is done with 2 up to 8 taxis and passengers.

The stable assignments are not the best possible, but this is also not expected. Given the fraction of the time, which the assignment creation needs, which is estimated to be only O(n^2), the results are more than acceptable, bearing in mind that they are still same or better than the first-come, first-served approach.

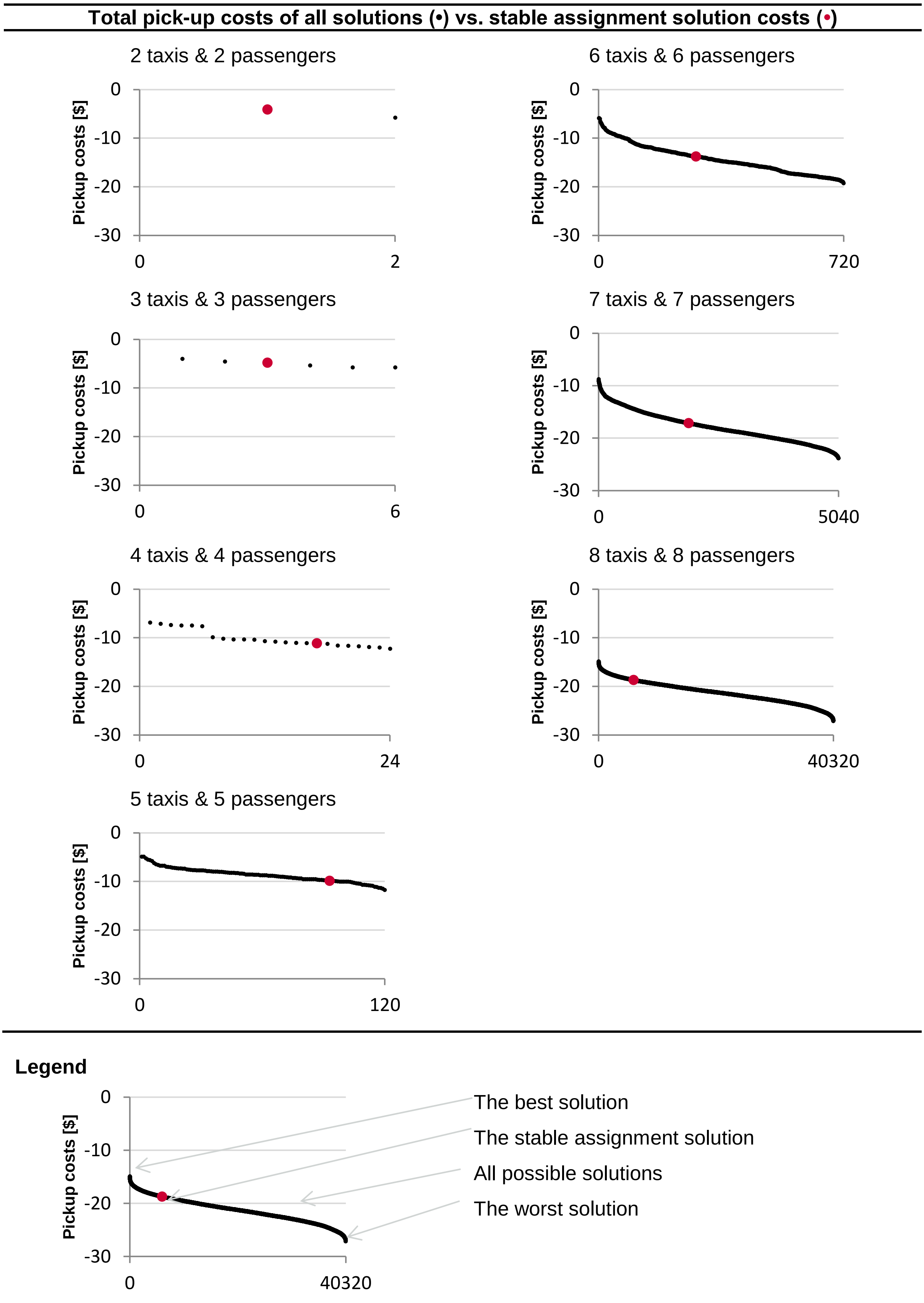


Figure 44 Price of stability estimation and comparison with all possible assignment solutions

4.2 Re-assignment of all confirmed immediate and advance bookings

This experiment pushes the application of the stable marriage algorithm even further to cover not just immediate booking requests but also advance ones, as described in chapter 3.1.4 “Re-assignment of all confirmed immediate and advance bookings”. In order to do so, the scheduling strategy first schedules all known new requests and then aims to re-schedule the plan in order to strive for the optimal assignment of taxis to passengers at all times. This section describes the experiment that was conducted and summarises the input parameters used, which appear in Tables 23 to 27.

Two versions of the stable marriage algorithm, where either taxis (SM_Taxis) or passengers (SM_Pax) propose first, are analysed. The strategies are analysed for 11 degrees of dynamism (DOD) ranging from situations where all booking requests are immediate (DOD = 100%) to those where all booking requests are made in advance (DOD = 0%). Moreover, three demand scenarios are examined. The demand level is defined by a ratio of booking requests per hour to taxis of either 3:1, 2:1 or 1:1, representing high, moderate and low demand scenarios respectively and corresponding to either 2400, 1600 or 800 passenger trips for 100 taxis over a period of 8 hours. The length of the decision epoch is set to 60s. (Kümmel et al., 2016c)

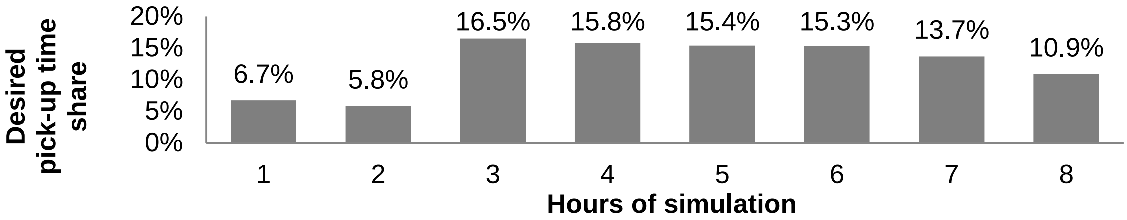


Figure 45 Distribution of the passengers’ desired pick-up times

The distribution of desired pick-up times is shown in Figure 45; it is designed to resemble the time period around a morning demand peak. The booking times of immediate booking requests equal to the beginning of the desired pick-up time interval. The booking times of advance booking requests are uniformly distributed between the beginning of the simulation and the beginning of their desired pick-up time frame. All simulations ended after 9 hours and 6 minutes of runtime to guarantee comparability of results. This time is set in order to enable all passengers to be delivered to their destinations.

The preference making is implemented similarly as in the methodology of dispatching taxis to immediate requests. The ties (= indifference in preferences), however, do have an impact on the strategy based on the stable marriage algorithm when the passengers propose first. In particular, advance booking requests can be accepted by many taxis and therefore the taxi with the lowest ID number is selected. This leads to higher utilization of less taxi vehicles.

The results are summarised in Figure 46 and Table 33. The best-performing strategy is highlighted in bold in each scenario. As expected, the FCFS dispatching strategy performs well in situations where all booking requests are immediate (i.e. where the degree of dynamism reaches 100%). As the share of advance booking requests becomes greater, the results show the increasing drawbacks of this simple assignment approach.

The number of served booking requests and the profit decline as the degree of dynamism approaches 0% (where all booking requests are made in advance). Moreover, the taxi 'on-call' distances and total distances rise. These trends hold true across high, moderate and low demand scenarios. Moreover, Table 33 reveals that the stable marriage algorithm performs best in reducing passengers' waiting time and total distance driven. It also improves the average taxi profit and the number of passengers picked up.

The high demand scenario uncovers the maximum performance capacity of the given taxi fleet size to serve the demand. As shown in the high demand section of Table 33, the maximum total distance driven across all experimental scenarios oscillates between 235 and 248 km per taxi, which indicates the saturated utilisation of taxis.

Figure 47 depicts the resulting schedules for 27 selected scenarios, involving all three strategies with the degree of dynamism set at 100%, 50% or 0% in high, moderate and low demand. Although the individual trip details cannot be discerned in these depictions, this figure provides a representation of fleet utilisation in the different conditions.

In the low demand scenario with DOD equal to 0%, the strategy SM_Pax requires 23 fewer taxi vehicles to serve the 800 advance booking requests, as illustrated by the red overlaid bar on the schedule at the bottom right of Figure 47. Using 77 instead of 100 taxis is saving a total of $1737 (23 taxis * 8.3 $/h * 9.1 h) and therefore the average taxi profit more than doubles from $27 to $57.

For example, the high demand scenario shows the transportation utilisation limits of the fleet under circumstances where only a few taxis are available, which is not the case for the scenarios of moderate or low demand. The low demand scenario reveals the property of the strategy SM_Pax to employ fewer vehicles than SM_Taxis and FCFS and to create chains of trips. Moreover, higher 'on-call' distances may be observed when the degree of dynamism is 0%. The taxis spent more time and distance reaching their customers, which in turn resulted in lower waiting time for passengers. (Kümmel et al., 2016c)

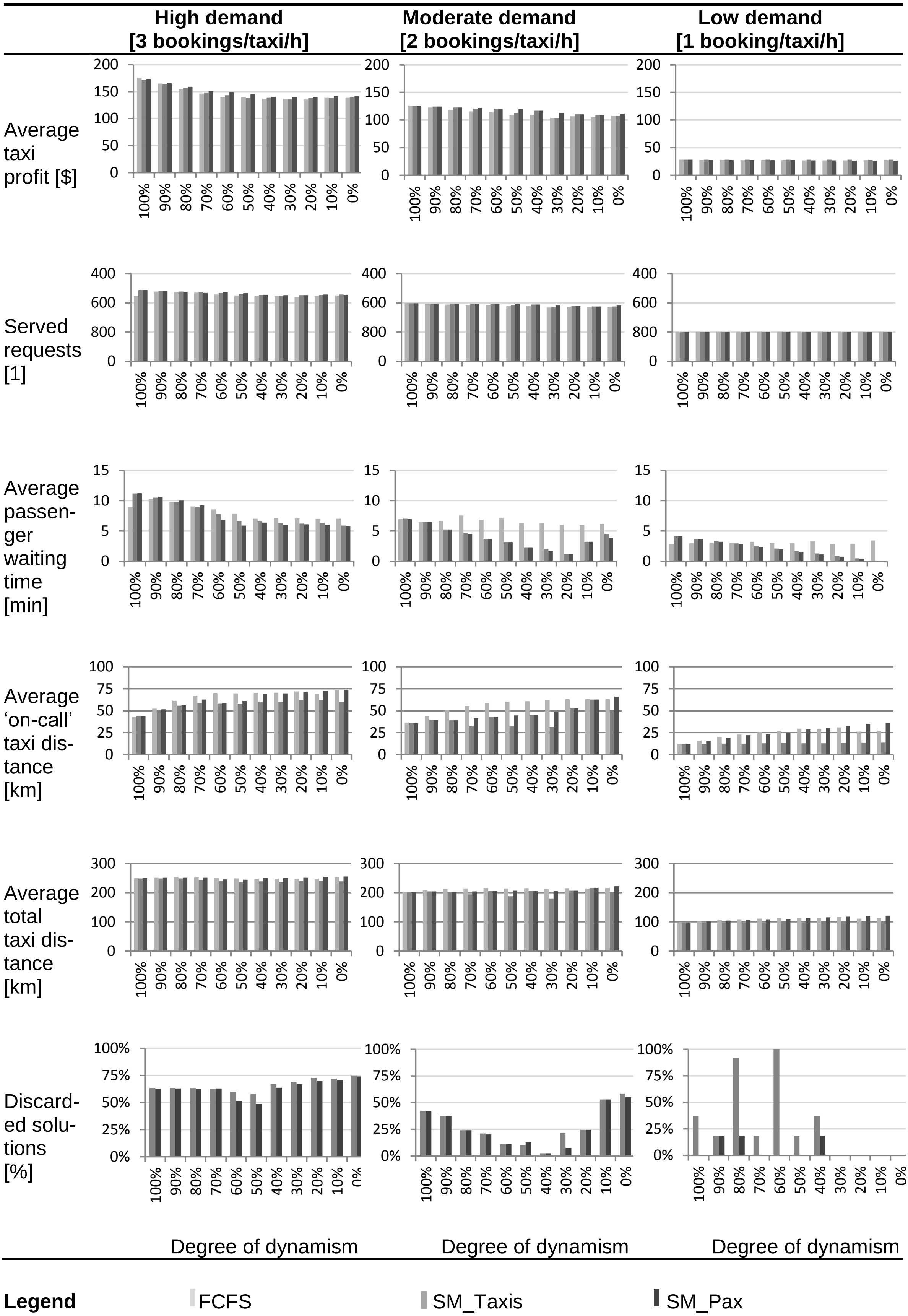


Figure 46 Performance of the compared strategies at DOD 100 to 0% (Kümmel et al., 2016c)

Demand	Performance	Strategy	100%	90%	80%	70%	60%	50%	40%	30%	20%	10%	0%
			Degree of dynamism										
High demand	Average taxi profit [$]	FCFS	**176**	165	154	146	140	139	137	137	135	139	138
		SM_Taxis	172	164	157	148	143	138	139	135	138	138	139
		SM_Pax	173	**166**	**159**	**151**	**149**	**145**	**141**	**140**	**140**	**142**	**142**
	Served bookings out of 2400 [1]	FCFS	1786	1906	1893	1877	1829	1802	1789	1791	1768	1795	1798
		SM_Taxis	**1952**	1931	**1909**	**1893**	1866	1839	1816	1793	1806	1816	**1827**
		SM_Pax	1950	**1937**	1901	1871	**1893**	**1857**	**1818**	**1808**	**1808**	**1828**	1823
	Average passenger waiting [min]	FCFS	**8.9**	**10.3**	**9.8**	9.0	8.6	7.8	7.0	7.2	7.1	7.0	7.0
		SM_Taxis	11.2	10.5	**9.8**	**8.9**	7.8	6.7	6.6	6.3	6.2	6.3	5.9
		SM_Pax	11.2	10.6	10.0	9.2	6.8	**5.9**	**6.4**	**6.1**	**6.1**	**6.0**	**5.8**
	Average 'on-call' taxi distance [km]	FCFS	**43**	52	61	67	70	70	70	71	72	69	73
		SM_Taxis	44	**50**	**56**	**58**	**58**	**58**	**60**	**60**	**62**	**62**	**60**
		SM_Pax	44	51	**56**	63	59	61	69	70	71	72	74
	Average total taxi distance [km]	FCFS	250	251	252	251	250	249	247	248	248	248	252
		SM_Taxis	**248**	**248**	**248**	**244**	**239**	**235**	**238**	**235**	**240**	**240**	**238**
		SM_Pax	249	251	251	251	245	244	249	250	251	253	255
	Discarded solutions [%]	SM_Taxis	63	63	63	63	60	58	67	69	73	72	75
		SM_Pax	63	63	63	63	51	49	64	67	70	71	74
Moderate demand	Average taxi profit [$]	FCFS	**126**	123	119	116	114	109	110	104	107	106	107
		SM_Taxis	**126**	**124**	**123**	120	**120**	113	**117**	104	**110**	**109**	108
		SM_Pax	**126**	**124**	**123**	**122**	**120**	**120**	**117**	**113**	**110**	**109**	**112**
	Served bookings out of 1600 [1]	FCFS	**1590**	1574	1554	1538	1531	1506	1504	1474	1486	1481	1488
		SM_Taxis	1588	**1580**	**1571**	1562	**1564**	1524	**1549**	1476	**1507**	**1498**	1500
		SM_Pax	1588	**1580**	**1571**	**1567**	**1564**	**1560**	**1549**	**1524**	**1507**	**1498**	**1523**
	Average passenger waiting [min]	FCFS	7.0	**6.5**	6.7	7.6	6.9	7.2	6.3	6.3	6.1	6.0	6.2
		SM_Taxis	7.0	**6.5**	**5.3**	4.6	**3.7**	3.2	**2.3**	2.0	**1.3**	**3.2**	4.5
		SM_Pax	**6.9**	**6.5**	**5.3**	**4.5**	**3.7**	**3.1**	**2.3**	**1.7**	**1.3**	**3.2**	3.8
	Average 'on-call' taxi distance [km]	FCFS	**36**	44	51	55	58	60	61	62	63	63	63
		SM_Taxis	**36**	**39**	**39**	**32**	**43**	**32**	**45**	**31**	**53**	**63**	**51**
		SM_Pax	**36**	**39**	**39**	41	**43**	45	**45**	48	**53**	**63**	66
	Average total taxi distance [km]	FCFS	203	208	212	214	216	214	215	212	215	214	216
		SM_Taxis	**202**	**204**	**203**	**194**	**205**	**188**	**205**	**179**	**207**	**216**	**203**
		SM_Pax	**202**	**204**	**203**	204	**205**	206	**205**	205	**207**	**216**	221
	Discarded solutions [%]	SM_Taxis	42	37	24	21	11	10	3	22	25	53	58
		SM_Pax	42	37	24	20	11	13	3	8	25	53	55
Low demand	Average taxi profit [$]	FCFS	**28**	**28**	**28**	**28**	27	27	27	27	27	27	27
		SM_Taxis	**28**	**28**	**28**	28	28	28	28	28	28	28	28
		SM_Pax	**28**	**28**	**28**	**28**	28	27	27	27	27	27	27
	Served bookings out of 800 [1]	FCFS	**800**	**800**	**800**	799	**800**	**800**	**800**	**800**	**800**	**800**	**800**
		SM_Taxis	**800**	**800**	**800**	**800**	**800**	**800**	**800**	**800**	**800**	799	**800**
		SM_Pax	**800**	**800**	**800**	**800**	**800**	**800**	**800**	**800**	**800**	**800**	**800**
	Average passenger waiting [min]	FCFS	**2.9**	**3.0**	**3.0**	3.0	3.2	3.0	3.0	3.3	2.8	2.9	3.4
		SM_Taxis	4.1	3.7	3.4	2.9	2.5	2.1	1.7	1.3	0.9	0.5	0.1
		SM_Pax	4.1	3.7	3.2	**2.8**	**2.4**	**2.0**	**1.6**	**1.1**	**0.8**	**0.4**	**0.0**
	Average 'on-call' taxi distance [km]	FCFS	**12**	16	20	23	26	27	30	29	31	26	27
		SM_Taxis	**12**	**12**	**12**	**12**	**13**	**13**	**13**	**13**	**13**	**13**	**13**
		SM_Pax	**12**	16	19	22	23	25	29	30	33	35	36
	Average total taxi distance [km]	FCFS	**98**	101	105	108	111	112	115	115	116	111	113
		SM_Taxis	**98**	**97**	**98**	**98**	**98**	**98**	**98**	**98**	**98**	**98**	**99**
		SM_Pax	**98**	101	104	107	108	110	114	115	118	120	121
	Discarded solutions [%]	SM_Taxis	0	0	1	0	3	0	0	0	0	0	0
		SM_Pax	0	0	0	0	0	0	0	0	0	0	0

Table 33 Performance of strategies at various DOD (100% to 0%) (Kümmel et al., 2016c)

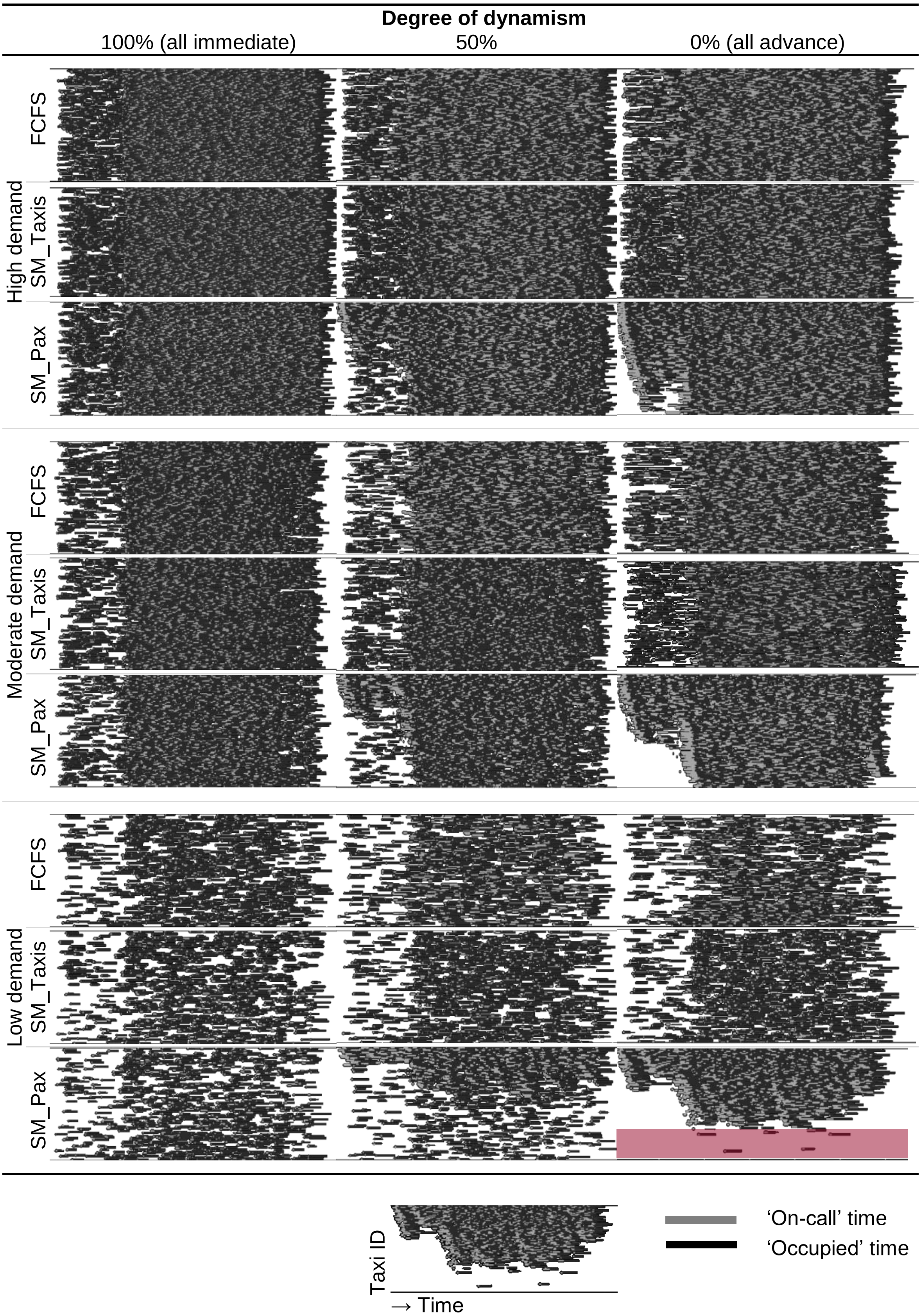


Figure 47 Resulting schedules for the 100%, 50% and 0% (Kümmel et al., 2016c)

If the results are further aggregated and the percentual improvements calculated, it can be observed, how the strategies SM_Taxis and SM_Pax perform in comparison to the FCFS strategy. Table 34 and Table 35 show a disadvantage of both strategies if the degree of dynamism equals to 1, in which the initial delay in response to the passenger results in worsening the average passenger waiting time. On the other hand, when DOD is approaching 0% both algorithms show their maximum improvement potential. Both strategies based on the stable marriage algorithm outperform the FCFS strategy in almost all observed performance indicators. Most notably, the strategy SM_Pax is able to serve more passengers and reduce their waiting times even more than SM_Taxis. The preferences of passengers and taxis are traceable in the results: Passenger's preference for the taxi that can arrive the soonest reduces the average passenger waiting time.

SM_Taxis Improvement	**DOD**										
	100%	90%	80%	70%	60%	50%	40%	30%	20%	10%	0%
Average taxi profit	-0.9%	0.6%	2.1%	2.5%	3.7%	2.0%	4.3%	0.9%	3.2%	1.4%	1.4%
No. of served bookings	3.1%	0.6%	0.6%	0.8%	1.4%	1.1%	1.5%	0.1%	1.2%	0.7%	0.8%
Average passenger waiting time	-23.8%	-8.8%	2.9%	14.3%	25.5%	34.0%	36.9%	46.6%	53.7%	46.2%	46.8%
Average 'on-call' taxi distance	-1.0%	13.1%	23.5%	33.2%	31.4%	38.6%	32.3%	40.0%	29.3%	20.1%	29.5%
Average total taxi distance	0.3%	2.2%	4.4%	7.4%	6.9%	10.1%	7.7%	11.5%	7.4%	4.7%	7.8%

Table 34 SM_Taxis improvement over FCFS (average of all demand levels)

SM_Pax Improvement	**DOD**										
	100%	90%	80%	70%	60%	50%	40%	30%	20%	10%	0%
Average taxi profit	-0.8%	0.5%	2.0%	2.7%	4.3%	4.7%	3.4%	3.5%	2.0%	0.7%	1.3%
No. of served bookings	3.0%	0.7%	0.5%	0.6%	1.9%	2.2%	1.5%	1.4%	1.2%	1.0%	1.2%
Average passenger waiting time	-23.2%	-8.9%	3.8%	15.0%	30.4%	38.7%	40.1%	51.1%	55.4%	48.4%	51.9%
Average 'on-call' taxi distance	-0.3%	4.7%	12.1%	11.8%	17.8%	15.7%	10.3%	6.8%	3.9%	-12.6%	-12.3%
Average total taxi distance	0.3%	0.7%	1.9%	1.9%	3.1%	2.5%	1.6%	0.6%	0.3%	-3.7%	-3.8%

Table 35 SM_Pax improvement over FCFS (average of all demand levels)

The in-depth insight into the results helps to understand how the proposed strategy performs during the simulation run under various demand levels through Table 36. One interesting insight with respect to demand levels provides Figure 48, which depicts the amount of taxis, which are 'vacant' during the simulation experiment. Here the strategy SM_Taxis is visualized in the high, moderate and low demand levels for both the two extreme cases: all bookings are immediate (DOD=100%) and all bookings are advance (DOD=0%). The impact of the demand is clearly observable in the number of taxis used. Whereas at low demand at least almost one half of taxis are 'vacant' during the whole simulation time, already at the moderate demand, almost all taxis are used. The red bar visualizes the minimal amount of taxis needed. The high demand scenario illustrates the fundamental performance limitations of the fleet to transport the demand.

Figure 48 also reveals the impact of advance bookings on taxi performance. The longer interval between the moment of booking and the desired pick-up time allows more vehicles to pick up the passenger and therefore to assign more appropriate one.

	Performance	Strategy	Degree of dynamism										
			100%	90%	80%	70%	60%	50%	40%	30%	20%	10%	0%
High demand	Average taxi profit	SM_Taxis	-2%	0%	2%	1%	2%	-1%	1%	-1%	2%	0%	0%
		SM_Pax	-2%	0%	3%	3%	7%	4%	3%	2%	4%	2%	2%
	No. of served bookings	SM_Taxis	9%	1%	1%	1%	2%	2%	2%	0%	2%	1%	2%
		SM_Pax	9%	2%	0%	0%	3%	3%	2%	1%	2%	2%	1%
	Average passen-ger waiting time	SM_Taxis	-25%	-2%	0%	1%	9%	15%	5%	12%	12%	10%	16%
		SM_Pax	-26%	-4%	-2%	-2%	20%	25%	9%	15%	14%	14%	18%
	Average 'on-call' taxi distance	SM_Taxis	-4%	5%	9%	13%	17%	17%	14%	15%	14%	10%	18%
		SM_Pax	-3%	2%	8%	6%	16%	12%	2%	1%	1%	-4%	-1%
	Average total taxi distance	SM_Taxis	1%	1%	2%	3%	4%	5%	4%	5%	3%	3%	5%
		SM_Pax	0%	0%	1%	0%	2%	2%	-1%	-1%	-1%	-2%	-1%
Moderate demand	Average taxi profit	SM_Taxis	0%	1%	3%	4%	6%	4%	7%	0%	3%	3%	1%
		SM_Pax	0%	1%	3%	5%	6%	10%	7%	9%	3%	3%	4%
	No. of served bookings	SM_Taxis	0%	0%	1%	2%	2%	1%	3%	0%	1%	1%	1%
		SM_Pax	0%	0%	1%	2%	2%	4%	3%	3%	1%	1%	2%
	Average passen-ger waiting time	SM_Taxis	-1%	0%	21%	39%	46%	56%	64%	68%	79%	46%	26%
		SM_Pax	0%	0%	21%	40%	46%	56%	64%	73%	79%	46%	38%
	Average 'on-call' taxi distance	SM_Taxis	1%	11%	23%	41%	26%	47%	26%	49%	17%	1%	20%
		SM_Pax	2%	11%	23%	25%	26%	26%	26%	22%	17%	1%	-4%
	Average total taxi distance	SM_Taxis	0%	2%	4%	9%	5%	12%	5%	15%	4%	-1%	6%
		SM_Pax	1%	2%	4%	5%	5%	4%	5%	3%	4%	-1%	-3%
Low demand	Average taxi profit	SM_Taxis	0%	1%	2%	2%	3%	3%	4%	4%	4%	2%	3%
		SM_Pax	0%	0%	0%	0%	0%	0%	0%	-1%	-1%	-3%	-3%
	No. of served bookings	SM_Taxis	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
		SM_Pax	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	Average passen-ger waiting time	SM_Taxis	-45%	-25%	-13%	3%	22%	31%	42%	60%	70%	83%	98%
		SM_Pax	-44%	-23%	-8%	7%	25%	35%	48%	65%	74%	85%	100%
	Average 'on-call' taxi distance	SM_Taxis	-1%	24%	38%	46%	51%	52%	57%	56%	58%	49%	50%
		SM_Pax	0%	2%	5%	4%	11%	9%	3%	-3%	-6%	-35%	-32%
	Average total taxi distance	SM_Taxis	0%	4%	7%	10%	12%	13%	15%	14%	15%	12%	12%
		SM_Pax	0%	0%	1%	1%	2%	2%	1%	-1%	-2%	-8%	-8%

Table 36 Detailed improvement over FCFS

Figure 49 provides an in-depth look at this phenomenon. Here the strategy SM_Taxis is visualized in the high, moderate and low demand levels and immediate and all advance booking situations are depicted. The number of taxis that can pick up new booking request is visualized for each decision epoch by a small dot. A moving average helps to see the trends. In accordance with the previous observations, Figure 49 reveals that the more advance bookings the more taxis are able to pick up the booking request.

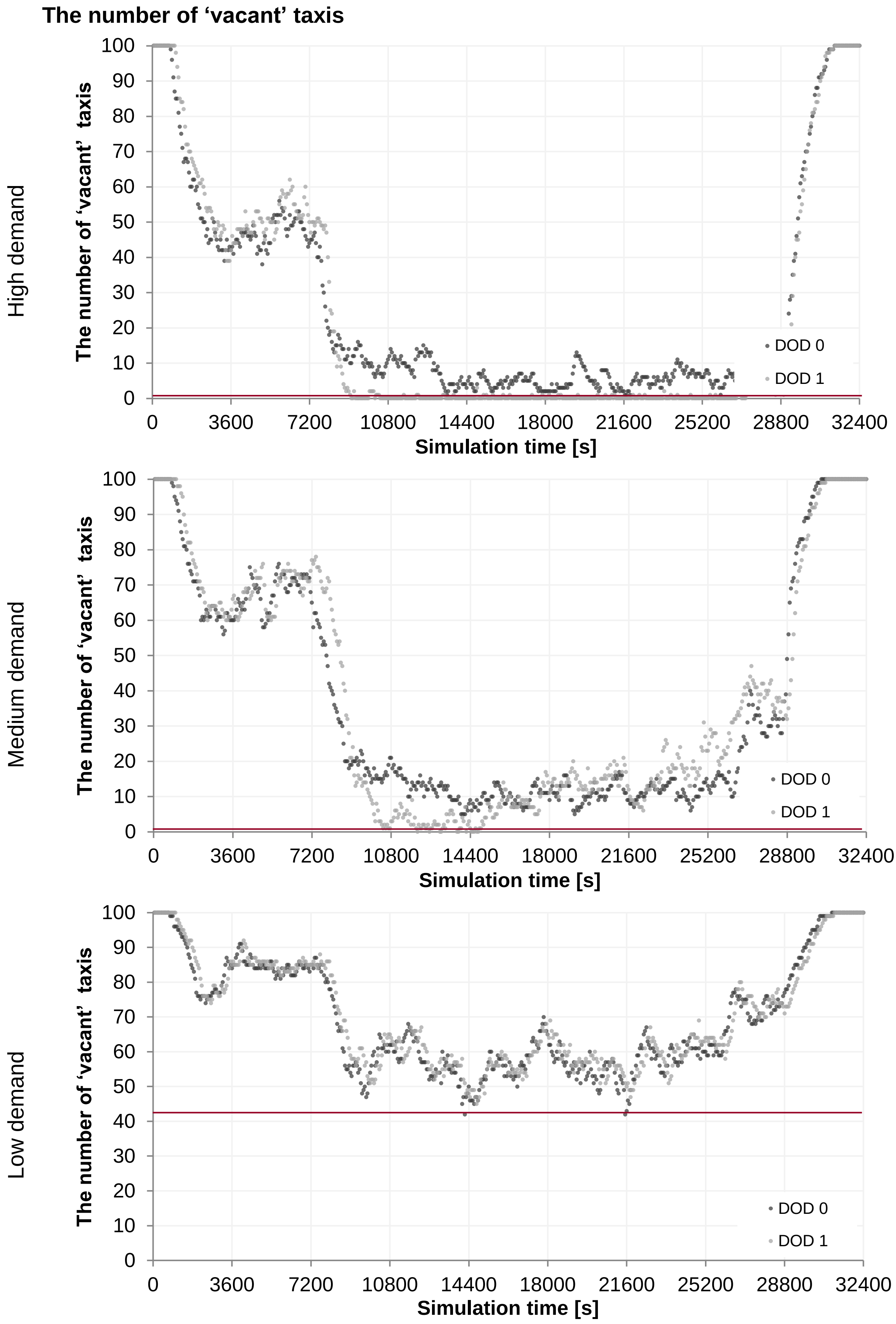


Figure 48 The number of 'vacant' taxis of SM_Taxis

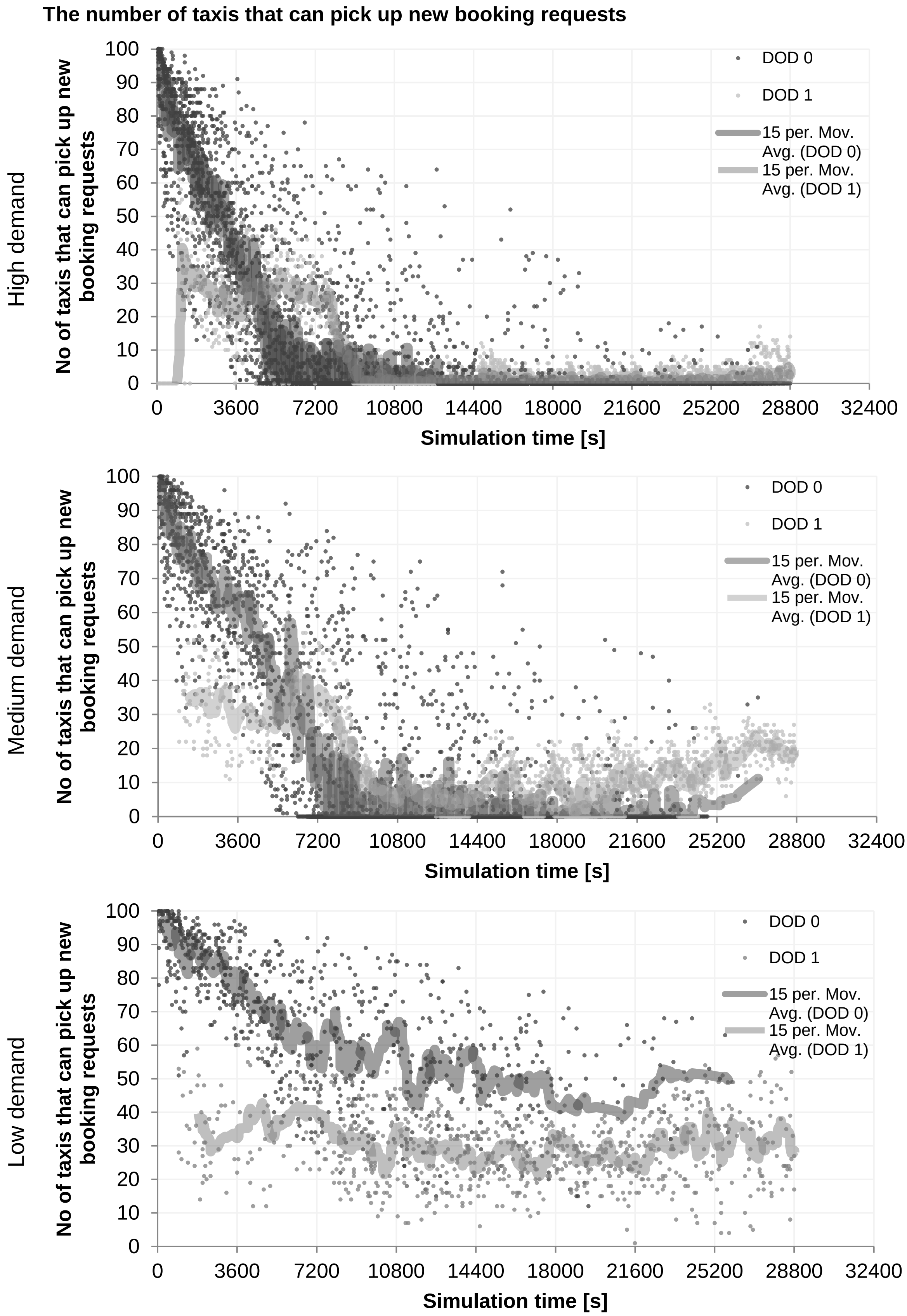


Figure 49 The number of taxis that can pick up new booking requests of SM_Taxis

The proposed methodology assigns and re-assigns the booking requests according to the stable marriage assignment algorithm that allows expressing preferences of not just the taxis but also the passengers. Moreover, the assignment can run in polynomial time and guarantees stability thanks to Nash equilibrium assignment.

In summary, the proposed re-assignment methodology based on constructive group re-assignment and the stable marriage assignment algorithm improves overall performance of taxi operations in comparison to the first-come, first-served method. The performance improvement is consistent across any combination of immediate and advance bookings and under low, moderate and high demand. The main advantage stems from the simultaneous group assignment of booking requests, which reduces 'on-call' distance, total distance, costs and passengers' waiting time. Moreover, group assignment accommodates the interests of both passengers, who wish to get a taxi in the desired time without waiting, and taxis, who prefer picking up the passengers, who impose the least cost.

Therefore the hypotheses: "The more advance booking requests, the better the taxi scheduling performance." cannot be confirmed. Theoretically, perfect knowledge about most of the booking requests should lead to really good results, but this is not the case. Indeed, they are even worse than if all requests are immediate. Why this is so? The answer is that there is still a proportion of not known trips.

4.3 Re-assignment of the timely nearest immediate bookings

This experiment answers the question whether taxis should accept individually sub-optimal re-assignments for the sake of system-wide performance in the way it is proposed in the section 3.1.5. Re-assignment of the timely nearest confirmed immediate bookings. Moreover, this experiment is based on real data, therefore, the following paragraph explains how.

The taxi dataset from San Francisco from 2008, which is generated for the project Cabspotting, is selected to be used for this experiment. ("Exploratorium Cabspotting," 2016) The dataset to be found here: ("CRAWDAD A Community Resource for Archiving Wireless Data At Dartmouth," 2016). For this project the movements and statuses of more than 500 Yellow Cab taxi vehicles were recorded throughout the greater San Francisco Bay Area for a year.

The dataset from San Francisco has a following format. Each taxi in active service recorded its position and status approximately every minute. The position is specified by longitude and latitude coordinates, the status is coded as occupied or not by logical value 1 or 0 respectively and the timestamp is given in Unix format as shown in Table 37.

From the changes in the status the trip origins and destinations can be approximately located and the trip begin-time and its duration estimated – as shown on Table 37. Moreover, the availability of the taxis can be reconstructed too. If the taxi is broadcasting the information, the taxi is on shift, if not; taxi has a break or is offline. Hence from the first information after being offline, the shift begin and the corresponding locations can be reconstructed. If the time difference between two consecutive time stamps exceeded 30 minutes, the taxi is assumed to be offline during this time. The last information in a row is taken to reconstruct the shift end time. (Kümmel et al., 2016b)

Raw data				Deduced information	
Longitude	**Latitude**	**Occupied**	**Timestamp**	**Occupied trips**	**Shift**
-122.437	37.76075	1	1213034892	Trip begin time & Origin location	Shift time begin & Initial taxi location
-122.442	37.76047	1	1213034955		
-122.442	37.75686	1	1213035023		
-122.443	37.75140	1	1213035056		
-122.443	37.74654	1	1213035057		
-122.444	37.74501	1	1213035102	Trip duration & Destination location	
-122.444	37.7451	0	1213035162		
-122.444	37.74607	0	1213035223		
-122.445	37.74929	0	1213035286		
-122.445	37.75248	0	1213035343		
-122.445	37.75298	0	1213035404		
-122.445	37.75309	0	1213035433		
-122.445	37.75283	0	1213035491		
-122.445	37.75283	1	1213035548	Trip begin time & Origin location	
-122.446	37.75049	1	1213035612		
-122.444	37.74834	1	1213035668	Trip duration & Destination location	
-122.442	37.75653	0	1213035729		Shift time end

Table 37 Raw data and deduced information

Figure 50 provides an illustration of the reconstructed trajectory from Table 37. There are two passenger trips and in between taxi moving unoccupied and searching for passengers.

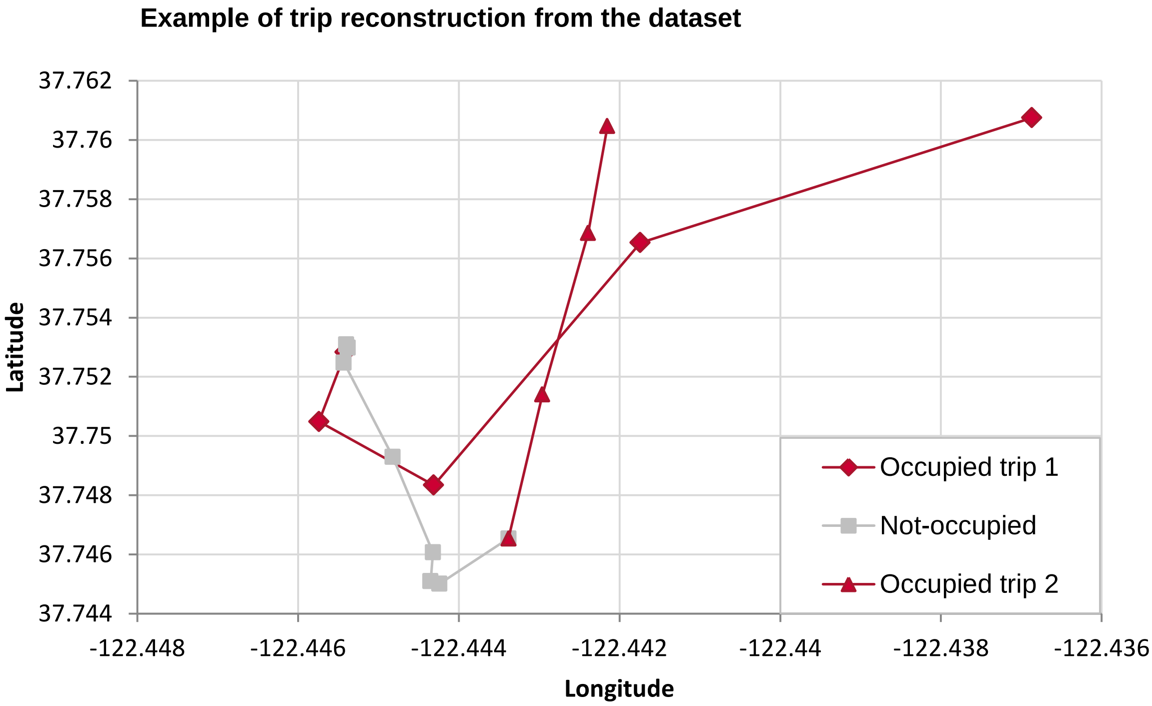


Figure 50 Illustration of the reconstructed trajectory from Table 37

Figure 51 and Figure 52 provide an illustration of all GPS coordinates recorded by two randomly selected taxis.

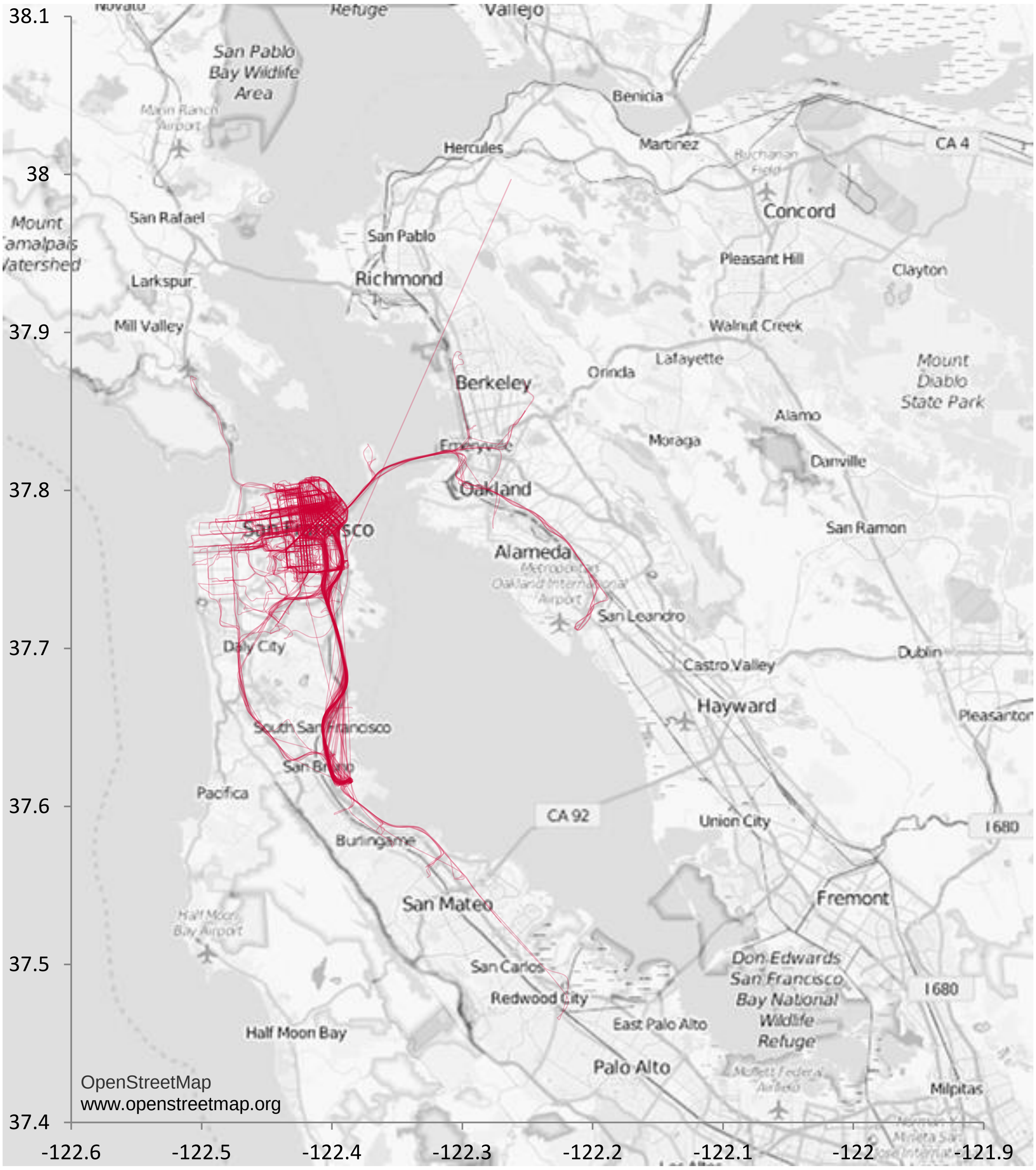


Figure 51 Randomly selected taxi records (1) (“OpenStreetMap,” 2016) (Kümmel et al., 2016b)

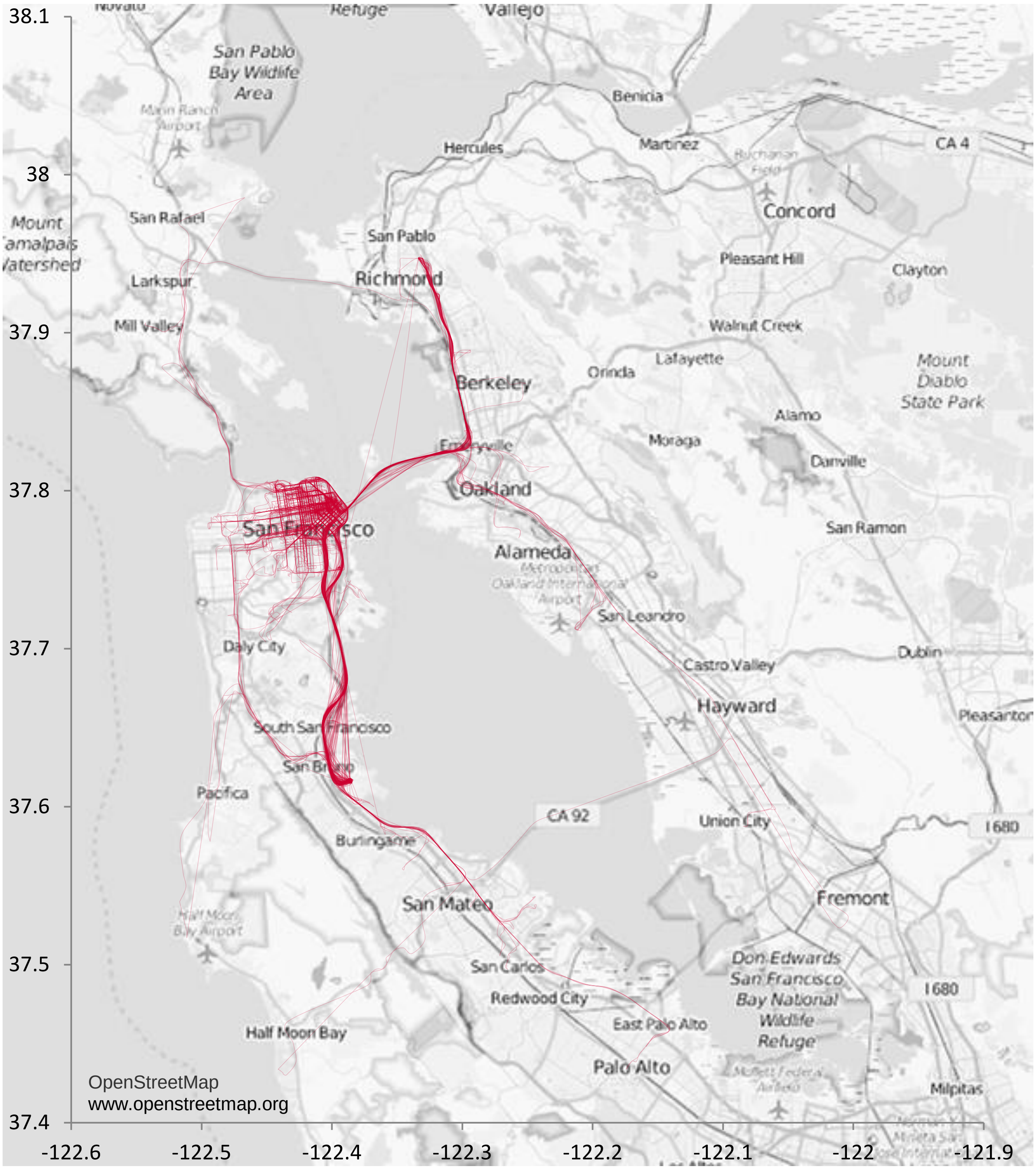


Figure 52 Randomly selected taxi records (2) ("OpenStreetMap," 2016) (Kümmel et al., 2016b)

With the above mentioned principles, the passenger trips and taxi shifts were reconstructed. For the purpose of the experiment, it is neither necessary nor practical to simulate the whole dataset length and therefore only an ordinary business day is selected: Monday 2008-06-02 06:00:00 – Tuesday 2008-06-03 06:00:00 (1212375600 – 1212462000 in Unix time).

It turned out that the raw data form the dataset contained some logical flaws most likely caused by wrong GPS position recognition. This is how they were handled. If a speed between two consecutive data points based on Euclidian distances exceeded 250 km/h the point is removed. Furthermore, all trips which duration is less than 30 s and distance of 0 were removed. Only trips, which started and ended in the above mentioned interval, were considered. As a result, Figure 53 illustrates the origin and destination locations of the 16302 considered trips. Figure 54 shows the distribution of time duration of these trips. Figure 55 shows the time distribution of trip time origins. (Kümmel et al., 2016b)

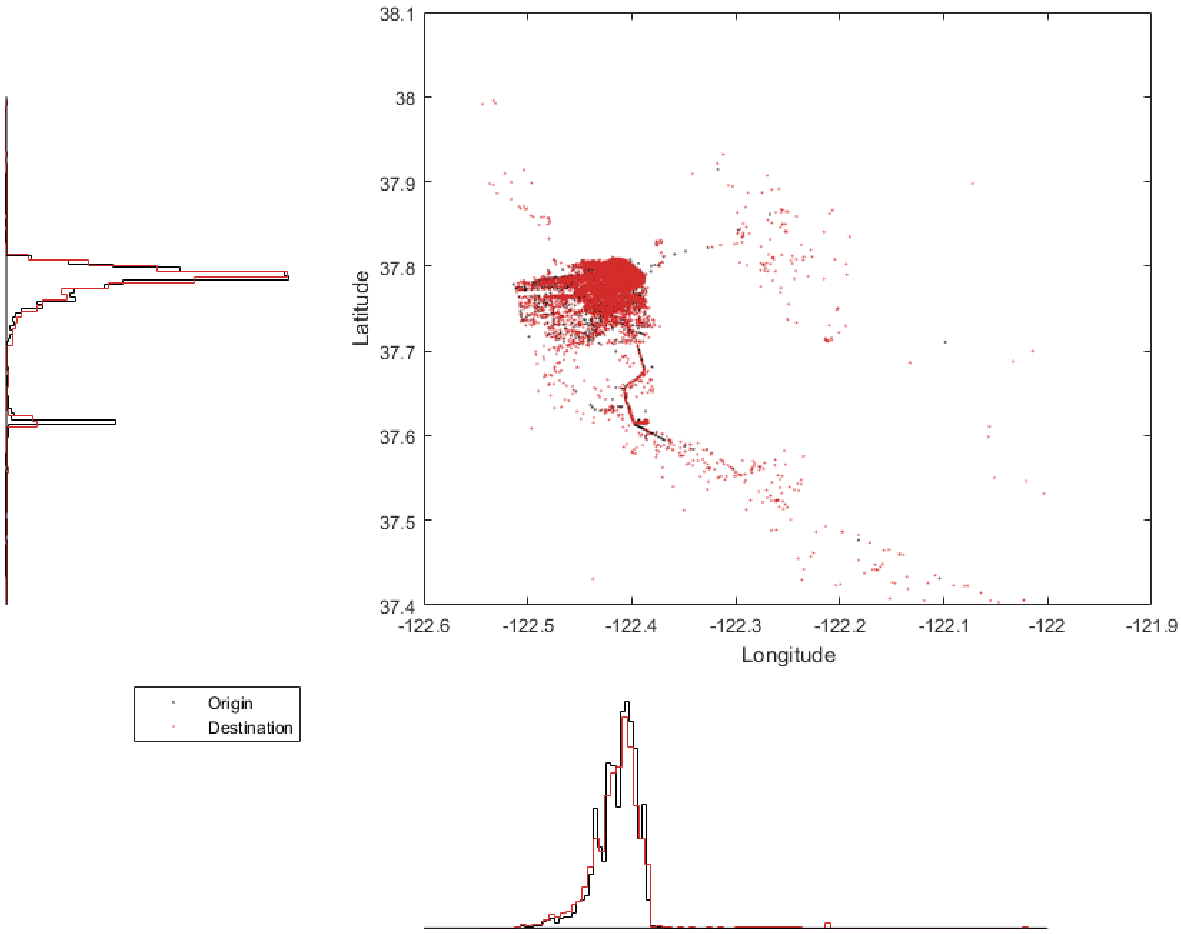


Figure 53 Origins and destinations of 16302 trips and their histograms (Kümmel et al., 2016b)

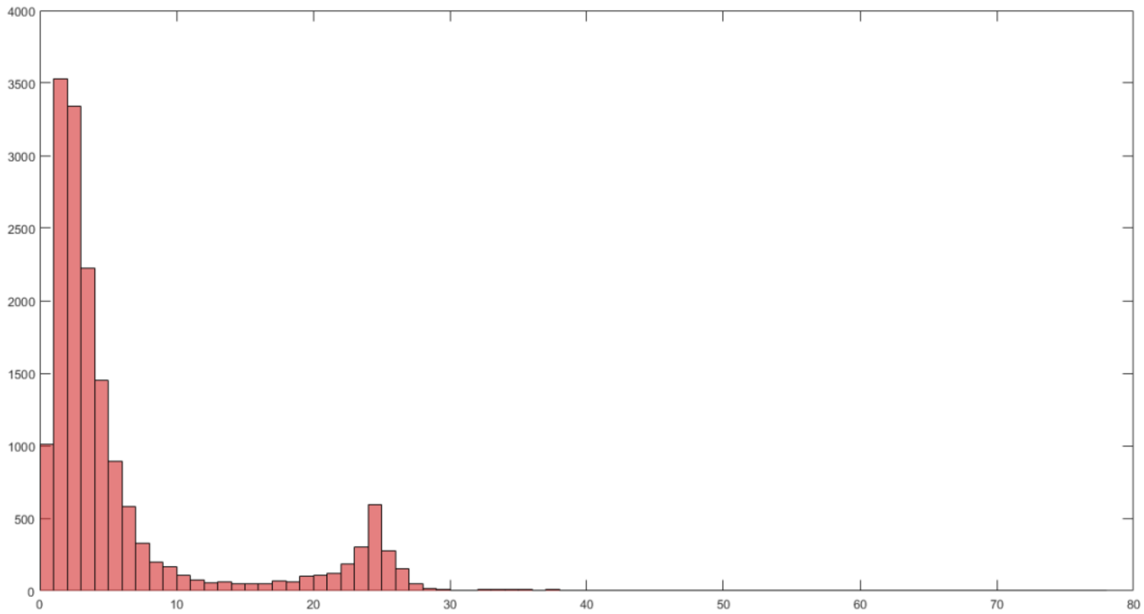


Figure 54 Distribution of time durations of 16302 trips in min (Kümmel et al., 2016b)

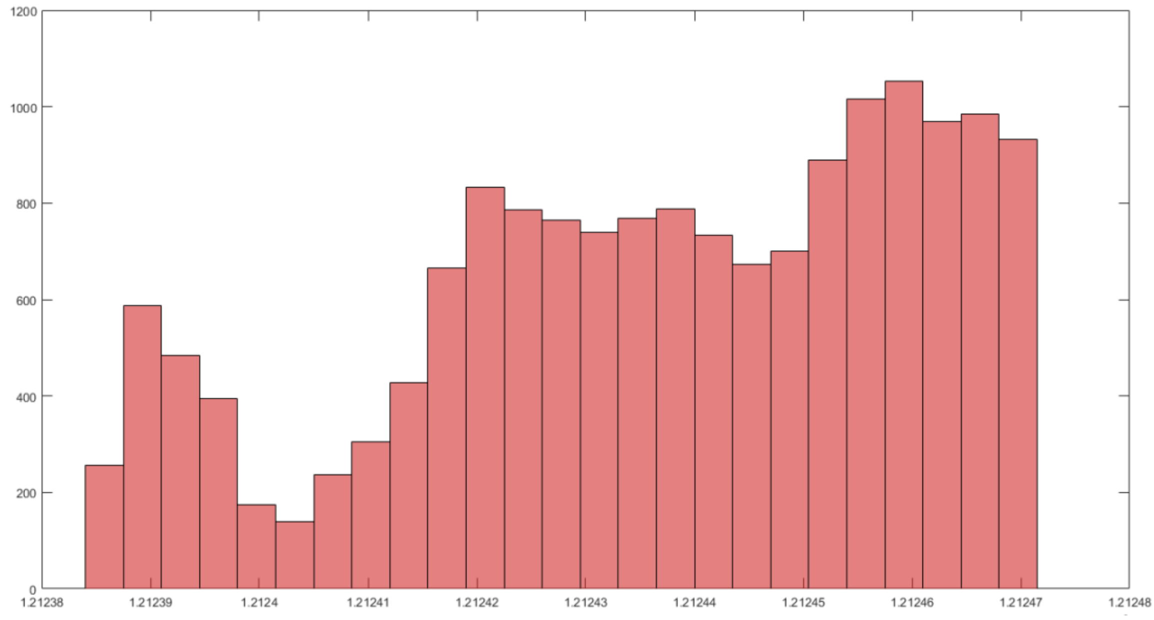


Figure 55 Pick-up times distribution 2008-06-02 06:00:00 – 06:00:00 (Kümmel et al., 2016b)

In order to get a better understanding of the experiment, similarly to the origins and destinations, it is also possible to plot the position of the taxis at their first shift – as shown in Figure 56. Moreover, Figure 57 summarizes the schedule of all 536 taxis, showing times, in which taxis were ‘occupied’ (red), ‘vacant’ (green) and ‘offline’ (black).

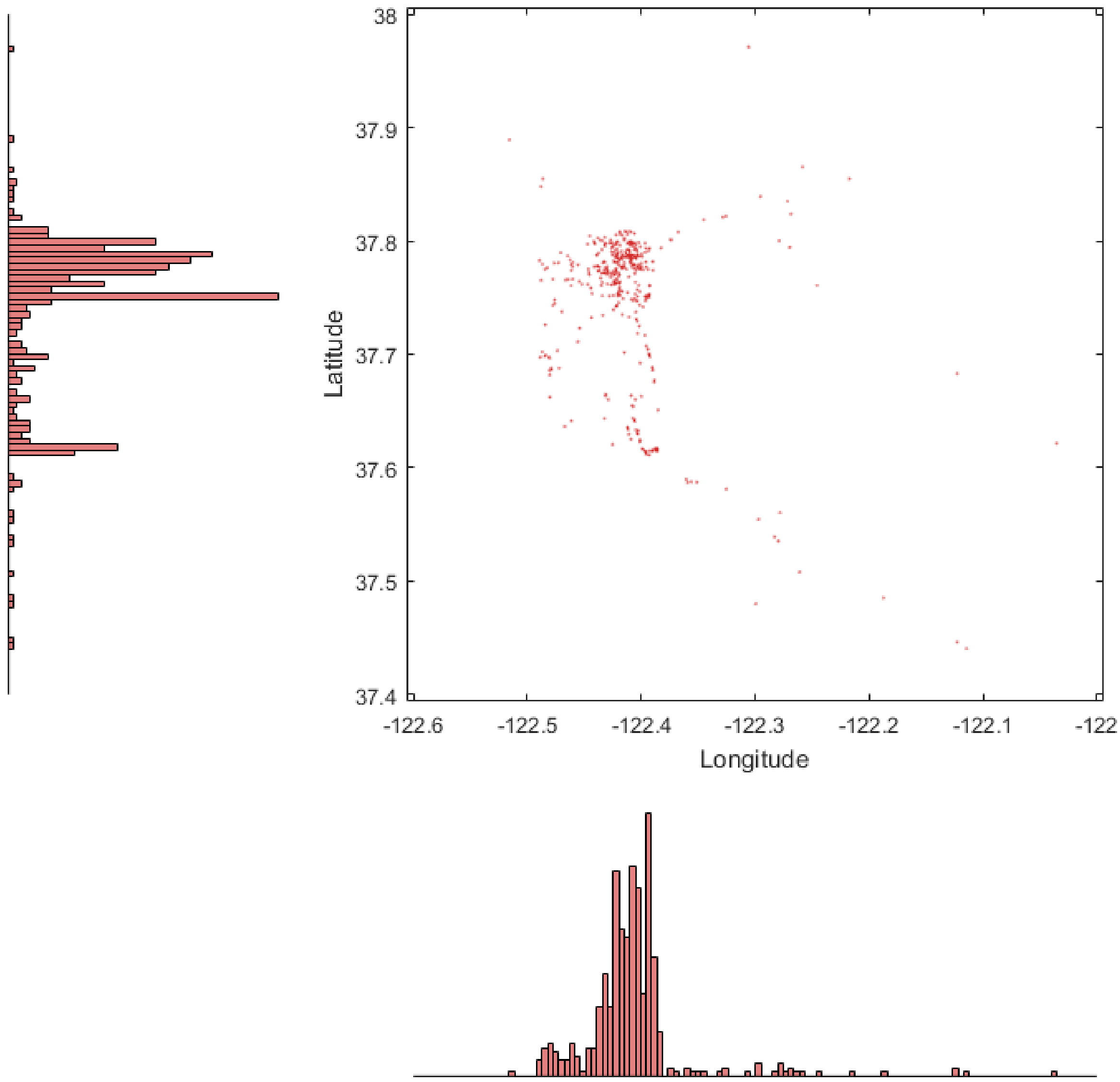


Figure 56 Initial taxi positons and their histograms (Kümmel et al., 2016b)

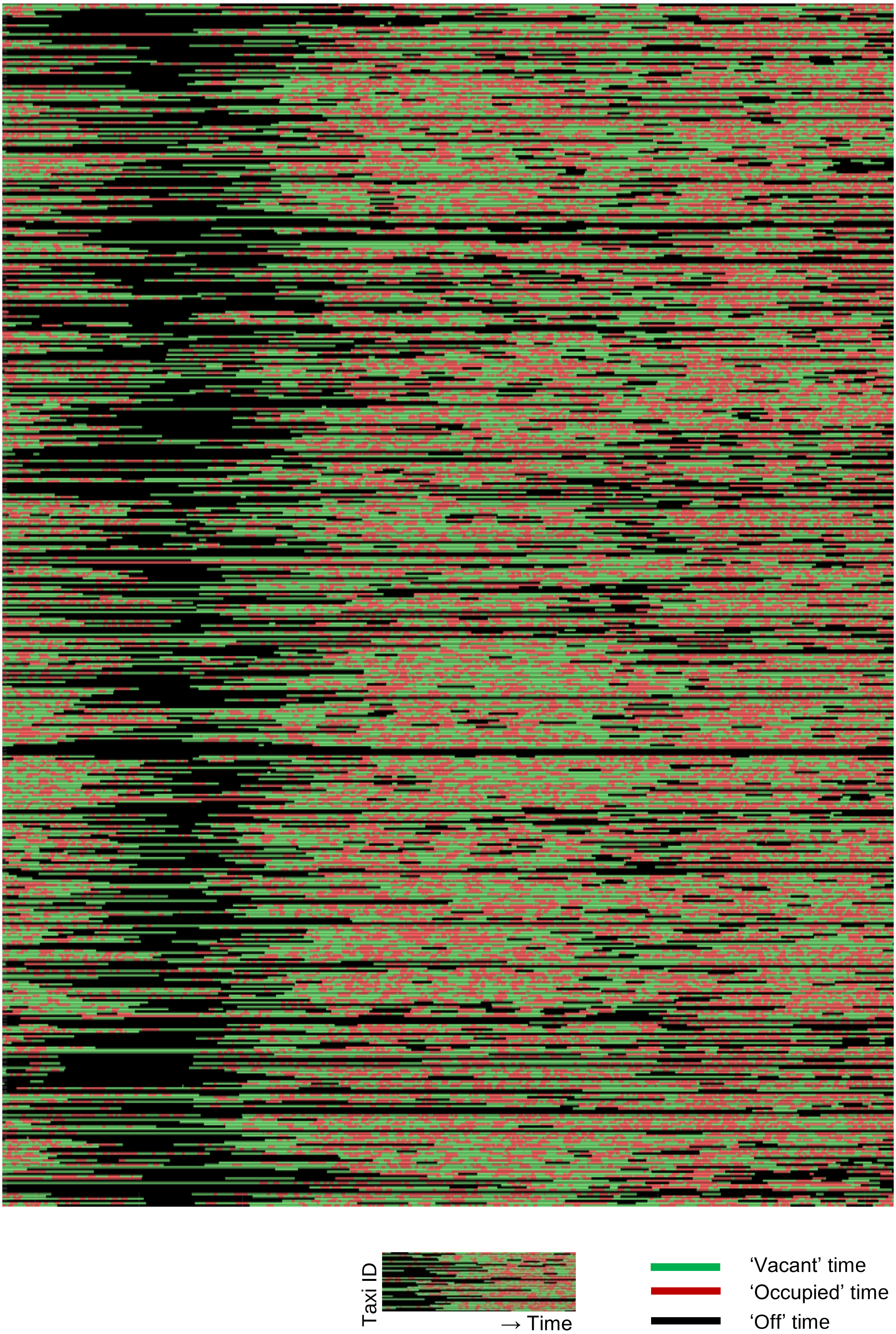


Figure 57 An illustration of schedules for all 536 taxis (Kümmel et al., 2016b)

The third experiment assumes all passenger trips are booked, drivers stay at their position after dropping off a passenger and once confirmed, taxis head to the next booking requests. At the simulation begin, all taxis are considered as ‘vacant’. Taxi no-shows, which happen to about a half of all confirmed taxi bookings in San Francisco, are not considered in the simulation (City and County of San Francisco Taxi Commission Public Convenience & Necessity Report, 2007).

In Experiment 3 three dispatching strategies are compared: the FCFS and two variants of the stable re-assignment based on Approach 3 called SM_Greedy_taxis and SM_Non-greedy_taxis, depending on the attitude of taxis to accept worse re-assignments than the current ones. The passengers are considered in both cases greedy, that is, unlike in Approach 2, the passengers only accept a re-assignment, which is same or better than the currently assigned taxi. For methodological details see chapter 3.1.5 “Re-assignment of the timely nearest confirmed immediate bookings”.

Table 38 provides an overview of the comparison of the two SM strategies to the FCFS strategy in absolute as well as relative terms. Since the ratio of demand to supply is relatively low, almost all passengers are picked up. The average ‘on-call’ and total distance has improved slightly. The highest improvement is recorded in the average passenger waiting time, which reached up to 5%. (Kümmel et al., 2016b)

	Absolute results			**Relative improvements**	
	FCFS	**SM_Greedy_taxis**	**SM_Non-greedy_taxis**	**SM_Greedy_taxis**	**SM_Non-greedy_taxis**
No. of served bookings [1]	16290	16290	16289	0.00%	0.00%
Average ‘on-call’ taxi distance [km]	38.86	38.73	38.74	0.34%	0.31%
Average total taxi distance [km]	186.21	186.09	186.08	0.07%	0.07%
Average passenger waiting time [min]	4.43	4.23	4.21	4.33%	4.92%
Combined improvement				**4.74%**	**5.30%**

Table 38 Comparison of the two SM strategies to the FCFS strategy (Kümmel et al., 2016b)

Tables 39 and 40 provide an overview about a sensitivity analysis results done on the main results. The length of the re-assignment (re-optimization) interval is varied on 1, 2, 3.75, 5, 7.5, 15, 20, 30 and 60 s for both Greedy and Non-greedy taxis.

	Re-assignment interval length (s)	**1**	**2**	**3.75**	**5**	**7.5**	**15**	**20**	**30**	**60**	**Avg.**
SM_Greedy_taxis	No. of served bookings [1]	16290	16290	16291	16291	16289	16290	16290	16290	16290	**16290.1**
	Average ‘on-call’ taxi distance [km]	38.62	38.88	38.97	38.77	38.46	38.82	38.68	38.68	38.68	**38.7**
	Average total taxi distance [km]	185.97	186.26	186.36	186.16	185.78	186.18	186.03	186.03	186.03	**186.1**
	Average passenger waiting time [min]	4.25	4.26	4.40	4.19	4.15	4.19	4.22	4.22	4.22	**4.2**
Improvement over FCFS	No. of served bookings	0.00%	0.00%	0.01%	0.01%	-0.01%	0.00%	0.00%	0.00%	0.00%	**0.00%**
	Average ‘on-call’ taxi distance	0.63%	-0.05%	-0.28%	0.23%	1.04%	0.11%	0.47%	0.47%	0.47%	**0.34%**
	Average total taxi distance	0.13%	-0.02%	-0.08%	0.03%	0.24%	0.02%	0.10%	0.10%	0.10%	**0.07%**
	Average passenger waiting time	3.93%	3.80%	0.63%	5.26%	6.33%	5.31%	4.56%	4.56%	4.56%	**4.33%**
	Combined improvement	4.69%	3.72%	0.27%	5.52%	7.59%	5.44%	5.13%	5.13%	5.13%	**4.74%**

Table 39 Detailed results of Experiment 3 for SM_Greedy_taxis dispatching strategy

	Re-assignment interval length	**1**	**2**	**3.75**	**5**	**7.5**	**15**	**20**	**30**	**60**	**Avg.**
SM_Non-greedy_taxis	No. of served bookings [1]	16289	16290	16290	16289	16289	16290	16290	16290	16290	**16289.7**
	Average ‘on-call’ taxi distance [km]	38.72	38.97	38.89	38.94	38.49	38.64	38.67	38.67	38.67	**38.7**
	Average total taxi distance [km]	186.04	186.32	186.25	186.26	185.80	185.99	186.03	186.03	186.03	**186.1**
	Average passenger waiting time [min]	4.34	4.26	4.26	4.21	4.11	4.13	4.19	4.19	4.19	**4.2**
Improvement over FCFS	No. of served bookings	-0.01%	0.00%	0.00%	-0.01%	-0.01%	0.00%	0.00%	0.00%	0.00%	**0.00%**
	Average ‘on-call’ taxi distance	0.37%	-0.28%	-0.08%	-0.20%	0.96%	0.57%	0.48%	0.48%	0.48%	**0.31%**
	Average total taxi distance	0.10%	-0.06%	-0.02%	-0.02%	0.22%	0.12%	0.10%	0.10%	0.10%	**0.07%**
	Average passenger waiting time	2.01%	3.82%	3.77%	4.79%	7.23%	6.76%	5.31%	5.31%	5.31%	**4.92%**
	Combined improvement	2.47%	3.48%	3.68%	4.56%	8.41%	7.45%	5.89%	5.89%	5.89%	**5.30%**

Table 40 Detailed results of Experiment 3 for SM_Non-greedy_taxis dispatching strategy

Figure 58 shows the combined improvement of both SM strategies under various lengths of the Re-assignment interval. The data stem from Tables 39 and 40. The figure suggests re-assignment interval longer than 20 s does not lead to any further improvements. On the other hand, an interval shorter than 5 seconds leads to volatile results. The volatility can be explained by the fact that the re-assignment interval may not cover some requests as the re-assignment is done every 5 s. Therefore, a ‘sweet spot’ with the highest improvement is around the re-assignment interval of 5 to 20 s.

Moreover, Figure 58 also helps to answer one of the main questions of Experiment 3. In the re-assignment interval 5 to 20 s the non-greedy strategy (SM_Non-greedy_taxis), the one under which taxis accept worse re-assignments than they currently have, outperforms the greedy strategy (SM_Greedy_taxis), in which the drivers accept only better assignments. (Kümmel et al., 2016b)

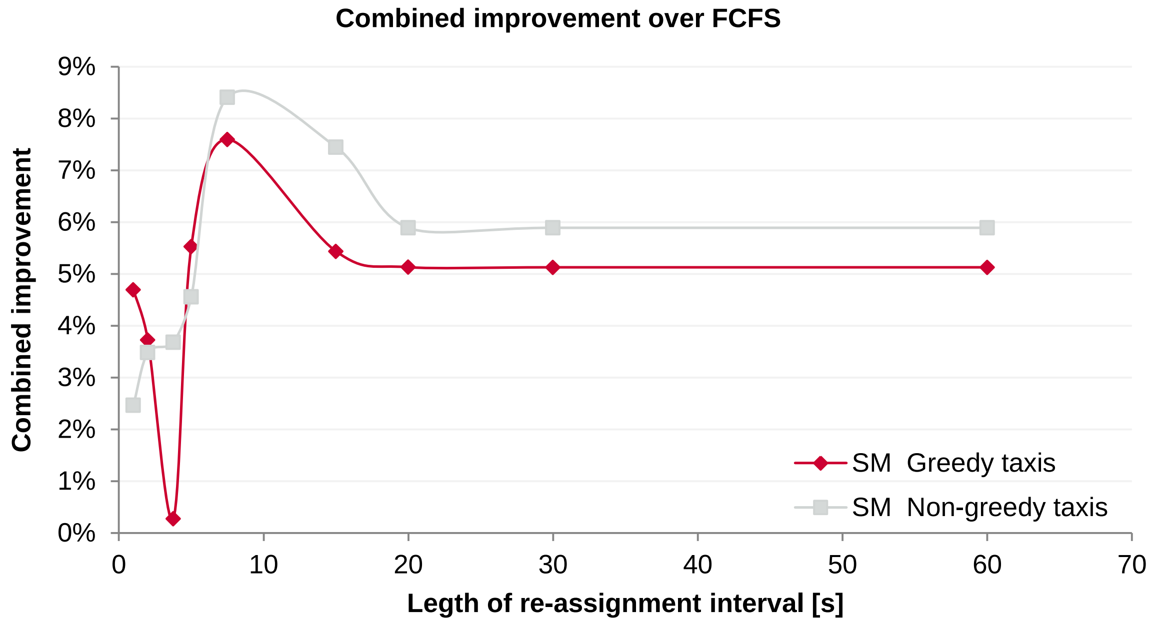


Figure 58 Combined improvement vs. the length of re-assignment interval (Kümmel et al., 2016b)

As a result, the hypothesis that: “Greediness of taxis does not pay off. Taxis should respect the system interest.” can be confirmed. The results show that the cooperative behaviour for the sake of system good pays off also on the individual level, at least in the average numbers.

5 Discussion and the lessons learned

The proposed methodology has been designed for a centralized system that assigns and re-assigns taxis to passenger booking requests simultaneously in real-time using the stable assignment. The results suggest that the proposed dispatching strategies improve most of the performance indicators across varying combinations of immediate and advance booking requests, degrees of dynamism and demand levels. Aside from the advantages and performance improvements, there are various trade-offs and fundamental limitations. This chapter discusses the broader limitations of the centralized approaches, buffering strategies, constructive re-assignment, stability considerations and implications of advance bookings. Moreover, it sketches a path to implement the proposed methodology in reality.

Centralized approaches

In any centralized approach, including the proposed one, there is a coordination entity (a scheduling and dispatching system) that manages the assignment of taxis to passengers. This entity has a complete overview of all bookings and taxi schedules. The coordination entity treats the vehicles as their ‘servants’ and requires drivers’ compliance with the assignments given to them.

From the taxi driver’s perspective, this might generate frustrations. (1) Firstly, it might happen, for example, that a driver ‘on-call’ to pick up one passenger is instead dispatched to get another passenger or told to wait. What is logical for the system might not seem logical for the individual. This frustration from frequent changes of planned tasks might be greatest at peak times. (2) Secondly, the drivers are required by the dispatching system to input their availability up front. The taxi’s availability is defined by times and positions at which its shift begins and ends. Within this time, the driver is subject to being dispatched as the system directs. (3) Thirdly, the system might require the individual taxi to act in such a way as to benefit the whole system but not the driver.

Decision epochs and buffering strategies

Decision epochs are time intervals used to buffer groups of booking requests in order to assign them simultaneously and to gain time for computing re-assignments of already confirmed bookings. What helps the overall performance of the taxi fleet has drawbacks for passengers. From the passenger’s perspective, a buffering strategy may prolong the initial waiting time to hear whether the booking is accepted or rejected. Moreover, this strategy requires a passenger to provide not just the desired origin and pick-up time but also the destination during the booking process. Whether the passengers will accept these trade-offs should be investigated in separate studies. Therefore, only if there are enough booking requests at the same time and at the same locations should buffering be used.

Constructive re-assignment of booking requests

Constructive re-assignment suffers from the fact that occasionally not all booking requests are re-assigned. Especially when the schedule is approaching its maximum limit in high demand scenarios (with ratios higher than two booking requests per taxi per hour), the procedure, unfortunately, ends up not scheduling some booking requests because the newly calculated schedule may break some of the pick-up interval constraints, as illustrated in Figure 59. In such a case, the results of the re-assignment must be discarded and the schedule before re-assignment should be used instead. The higher the proportion of discarded schedules, the more similar is the results to the FCFS scenario. This drawback applies, however, only to Experiments 2 and 3.

The motivation for using the constructive re-assignment of booking requests is the identified research need to overcome insufficient time for assignment calculation in the schedule adaptation approach. Hence, the constructive re-assignment, approach does not truly solve the problem. It offers an alternative solution which may be working very well and perhaps better, but whose quality similarly decreases at demand peaks. However, in order to draw definitive conclusions about this issue, the available experiments would not suffice. Comparing both approaches on a very large-scale city scenario would reveal the true capabilities of schedule adaptation and constructive re-assignment approaches.

Some booking requests cannot be re-assigned within their desired pick-up time

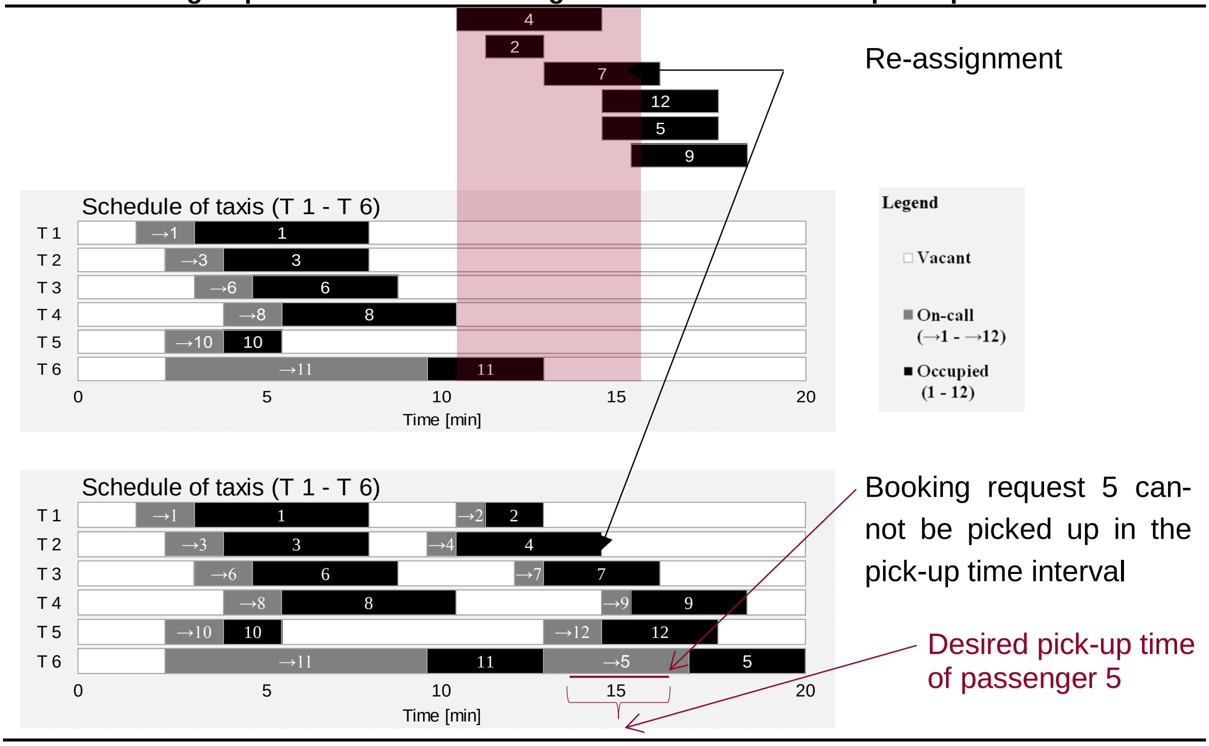


Figure 59 Some booking requests cannot be re-assigned within their desired pick-up time

Stable assignments of groups of booking trips to groups of taxis

The stable marriage assignment has difficulty handling an assignment of more trips to the same taxi, since the original algorithm is designed to assign one man to each woman (or vice versa). Figure 60 shows an example of the implications of this in a situation before and after re-assignment. Since the assignment insists on one-to-one matching, the group of six passenger requests is re-assigned to six taxis - one each. Unfortunately, for requests 11 this means a substantial prolongation of the 'on-call' time. Extending the stable marriage algorithm to allow matching multiple women to one man in time would be meaningful in future work.

Another limitation of the classical solution of the stable marriage problem is the simplification of the preferences to integer numbers describing the position in the preference list, but not necessarily the relation of how much worse or better the next partner is in the preference list. Removing this simplification may potentially further improve the results.

Finally, looking ahead to fully autonomous systems, where vehicles do what they are told and don't heed individual profit, the stability aspect will become superfluous and obsolete. Computers presently, unlike humans, don't mind how much they earn.

Fundamental limitation of the stable marriage algorithm in re-assigning bookings

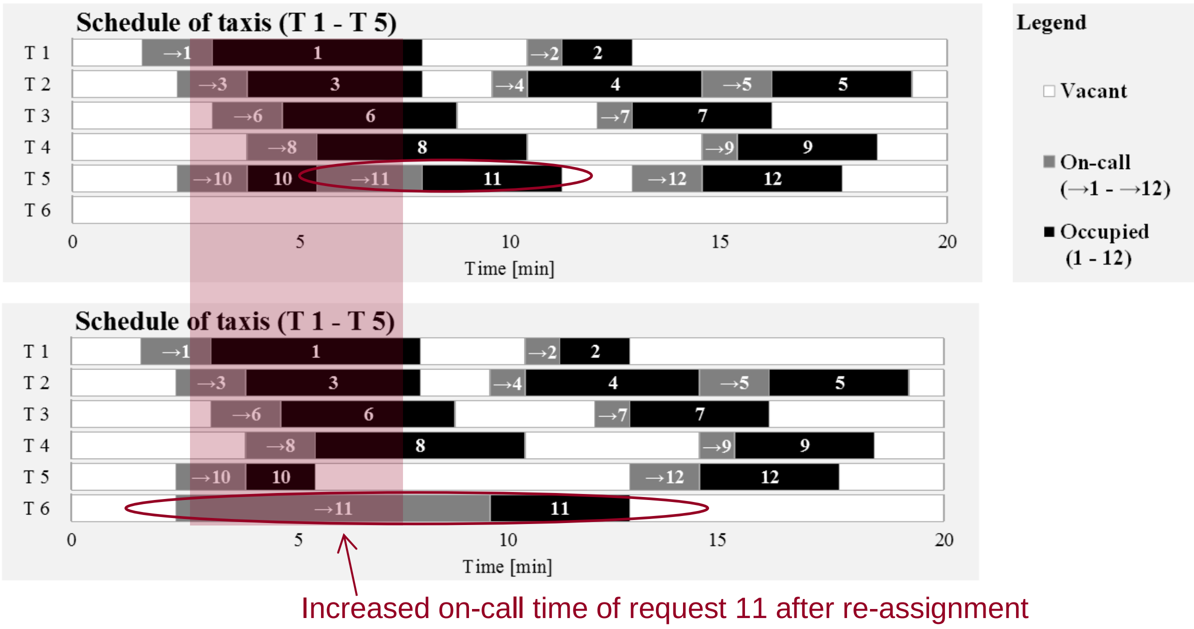


Figure 60 Fundamental limitation of the stable marriage algorithm in re-assigning bookings

Implications of advance bookings

With an increasing share of advance bookings the waiting time for taxi decreases, but so does the taxi profit (see Figure 46). Here are possible explanations.

Advance bookings (as defined in this study) allow passengers to book a taxi more in advance from 15 min to 24 h from the current moment. This means that the schedule is not fixed for the whole day ahead and the schedule cannot be calculated overnight (as it used to be with some dial-a-ride services). Fixed schedule would presumably lead to the best operational performance. This is not the case in taxi operation, however, because such service inflexibility is not acceptable for many passengers.

The proposed methodology, among others, aims to balance the interests of taxis to reduce their costs and passengers' interests to reduce waiting time. Perhaps another assignment methodology favouring only the taxi fleet interests and pursuing not stable but strictly optimal assignments would lead to better results.

The results in the high and moderate demand scenario may be significantly influenced by the number of discarded solutions (described above). The more discarded solutions, the more similar are the results to the sequential assignment strategy results.

In any case, if the taxi dispatching system allows also advance bookings, algorithms that re-assign some of the requests should be used in order to prevent worsening of the taxi schedules and corresponding passenger waiting times and taxi costs.

Simulation models as a tool for evaluating dispatching strategies and beyond

Two well-known quotes "All models are wrong, but some are useful" and "The practical question is how wrong do they have to be to not be useful" open a debate about the usefulness of the experimental results in verifying the proposed methodology. (Box and Draper, 1987) Simulation models are deemed to be the best known methods to estimate the potential of a new dispatching strategy. In particular, the agent based models, which allow modelling each taxi and each passenger coupled together with realistic routing and travelling time estimation as well as real-time dispatching strategies, in the future may not just serve to evaluate the benefits of a new dispatching strategy, but in fact serve as an integral part the dispatching strategies. Especially, for the fully autonomous fleets these models may serve as a communication backbone to gather and exchange information and facilitate decision making.

Path to implementation in reality

There are technical and business related steps in the implementation of the proposed approaches to reality. In a nutshell, the proposed methodology is designed for a centralized dispatching entity, which assigns the taxis to pick up passengers, in order to leverage on the efficiency gains by simultaneous group assignments. Since the assignment of individual taxis has a direct effect on the revenue earned by the individual taxi, it is not particularly suitable for the business model with self-employed drivers, but more for the one, which employs the drivers on a contract basis.

It is possible to either implement an assignment algorithm into an existing dispatching system instead of the conventional dispatching logic or to build a new dispatching solution around the algorithm and provide it as "software as a service". Depending on the existing system, the first approach may require some modifications. For example, the passenger interface may now ask also about the destination at the time of booking, may display the taxi ID first shortly before pickup, and show the countdown to the reply. The interface to taxi driver, should now ask the driver for his/her availability and communicate the assignments of his/her taxi to passenger bookings. Ideally, the system should be supporting the taxi driver in the decision making such as departure time selection, route choice and others. Ideally, the system should show the schedule of the taxi driver in full. The second option that aims to provide a dispatching system as a service would require more development. In particular, the speed of information transfer is crucial for the viability.

In both cases, the travel times between various points need to be estimated. The more precise the estimation will be, the better. Since the amount of estimations grows with O(n!) it makes sense to reasonably limit the travel time enquiries to the nearest ones or even better, to create a heuristic-like database capable of providing a quick and sufficiently precise answer to quickly preselect several assignments and then estimate the travel times of these assignments very precisely. The business model may either charge a one-time fee for using the service or be transaction-based. The height of this revenue may be either static, depending on the actual fare price, or derived as a share from the savings as compared to the previous dispatching logic.

Apart from these issues, the application should be extensively tested on the real data of the particular city but still in the simulation environment, in order to assess more precisely the improvement potential and in order to scrutinize the functionality and set and tune parameters to fit the particulate context and maximize benefits. Later, if proven to be successful, the system could be fully rolled out.

Commonly there are more than one taxi operators in a city and each has its own dispatching system. The proposed methodology can work also when applied to many operators in one city. However, the number of possible taxis may be reduced, because they do not belong to certain taxi operator. A meta-operator, which would virtually allow booking a taxi of any company, would help to improve the operation on the assumption that the passenger are not keen to go with a particular company.

Summary

To sum up, despite the above mentioned fundamental limitations in the methodology and its evaluation, the research of new methodologies for taxi dispatching is meaningful and potentially impactful.

6 Conclusions and future work

This thesis proposes a dispatching strategy for taxis to leverage on opportunities in taxi dispatching especially in large-scale taxi operations. The results indicate that proposed dispatching methodologies have the potential to serve more passengers with the same fleet of taxis, reduce passenger waiting time, and help operators and dispatchers to generate more profit.

The main advantage stems from the simultaneous stable assignment of groups of booking requests to a group of taxis that accommodates the interests of both passengers and taxis, which in turn reduces 'on-call' distance, total distance, costs and passengers' waiting time. The main disadvantage stems from the fact that the central coordinating entity requires drivers' compliance and the buffering strategies make passengers wait little longer for the initial reply.

The three proposed approaches can be clustered based on the degree of dynamism and the average demand to supply ratio. Approach 1 focused only on immediate booking requests (degree of dynamism DOD=100%) and is simulated in a scenario with 1 booking request per taxi and per hour. Approach 2 aimed at any combination of immediate and advance booking requests (DODϵ<0%,100%>). Approach 3 based on real data aimed at immediate requests and the average demand to supply ratio is around 1.3 booking requests per taxi per hour.

Table 41 shows all three experiments. All are able to run in real-time, where Experiment 1 and 3 are faster than Experiment 2 because of the absence of advance booking requests. Experiment 2, on the other hand, showed the highest improvement potential thanks to the simultaneous re-assignment of already confirmed requests in groups.

	Experiment 1	**Experiment 2**	**Experiment 3**
Advantages	+ Good results	+ Best results + Any combination of immediate and advance booking	+ Can cope with taxi breaks + Moderate improvement
Disadvantages	- No advance bookings	- Slowest	- No advance bookings

Table 41 Simplified advantages and disadvantages of the proposed three approaches

Table 42 summarizes the answers to the three hypotheses. In summary, it is found that the simultaneous assignment of new booking requests outperforms the sequential assignment if there are enough booking requests at similar location at similar time. This confirms the first hypothesis. It is expected that information about requests in advance allows the dispatching algorithm to outperform all the immediate booking requests. On contrary with the expectations, this was not proven in Experiment 2 and therefore the second hypothesis cannot be confirmed in this respect. The results showed clear benefits of simultaneous re-assignment of already confirmed booking requests; however, the dispatching with 100% of immediate requests still performs the best in the experiments made. Finally, the third experiment found out that greediness in the re-assignment process does not pay off and it makes sense for taxi drivers to accept the assignments with respect to the common system interest.

	Hypotheses evaluation	Result
1.	The potential of simultaneous assignment to improve taxi dispatching performance increases with the amount of booking requests that can be assigned at the same time and works the best for immediate bookings.	✓
2.	The more advance booking requests, the better the taxi scheduling performance.	✘
3.	Greediness of taxis does not pay off. Taxis should respect the system interest.	✓

Table 42 Evaluation of research questions and hypotheses

Simultaneous assignment of new booking requests as well as simultaneous re-assignment of already confirmed ones is beneficial for taxi performance. Introducing advance booking options for passengers without substantially changing the dispatching system is counter-productive and decreases the performance.

This work apart from the application of the stable marriage assignment algorithm for taxi scheduling and dispatching, also systemizes the method estimating the costs and profits of individual booking requests, handling the schedules, and also maps the fundamental limitations of not just the stable marriage algorithm but also the commonly shared aspects with other approaches.

It is the intention to provide initial and general insights into the performance of taxi booking with the stable marriage algorithm. The results are credible within the limits of the listed assumptions. They are promising and at the same time they reveal some fundamental limitations of the approach as discussed before. It cannot be excluded at this point that a very rigorous examination of the details of the simulation model and the assignment methodology may lead to further improvements of the performance of taxi operation.

Future work

With respect to the application in reality, the proposed strategies need to be further scrutinized for their robustness in various supply and demand levels in stochastic environments, in which some requests are cancelled, traffic slows down unexpectedly, vehicles break down, communication is delayed or lost, sensors are imprecise, location estimations or map-matching of position to road network is incorrect or not available, and the street topology or turn and stopping restrictions are not updated. In particular, learning capability and feedback loops can be built to help to deal with imprecisions and wrong estimations.

Moreover, the potential impacts of the mechanism design of the assignment based on the solution to the stable marriage problem should be examined in a separate study. For example, the results might be further improved by the non-greedy assignment of booking requests, a greedy scheduling system that accepts only better-performing schedules, an adjusted order of proposal modifications and adjustments in the method of selecting preferences. Especially the impact of combinations of varying many parameters at the same time should be investigated. In addition, it would be vital to include the dynamic fares and costs. Lastly, extending the stable marriage algorithm to allow matching multiple women to one man in time would be very meaningful.

A future study may survey in depth whether passengers are willing to trade shorter total waiting time (from the moment of booking until pick-up) for a longer wait for the initial reply from the dispatching system. In addition, it would be relevant to study whether taxi drivers are willing to give away some of their freedom in choosing, which passenger to pick up in exchange for increased earnings.

Real-time taxi dispatching is only one of the many challenges faced by taxi operators. Operators are also challenged to appropriately set the taxi vehicle fleet size to effectively meet the demand and to pre-allocate taxis in expectation of future demand. In order to apply dispatching systems holistically, the fleet sizing and pre-allocation mechanisms should be included. For automated vehicles, the planning of service, charging / refuelling and cleaning activities should be included as well. In terms of the stable marriage algorithm, the results might be further improved by adjusting the order of proposal modifications and modifications in the preference selection methodology. Further studies can also compare the benefits and trade-offs of all known dispatching strategies.

The proposed methodology and the whole framework of describing the problem may also find suitability in other related domains, especially where there is a need for a coordinated matching system in which two parties care about the outcome of the assignment. The candidate domain may range from warehouse and storing logistics, through the robotic manufacturing process, to the assignment of same-day urban delivery tasks.

Benchmarking the performance of the stable assignment with other assignment approaches would be worthwhile. Ideally, there should be two benchmarks: One simple benchmark, similar to the one used in this thesis, and the most elaborate one, based on real demands, supplies and network – perhaps based on the work by Maciejewski and Bischoff (2015) based on Berlin data.

References and citations

Aarhaug, J., Skollerud, K., 2014. Taxi: Different Solutions in Different Segments. Proceedings of 41st European Transport Conference Selected Proceedings 1, 276–283. doi:10.1016/j.trpro.2014.07.027

Afian, A., Odoni, A., Rus, D., 2015. Inferring Unmet Demand from Taxi Probe Data. IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC 2015-October, 861–868. doi:10.1109/ITSC.2015.145

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

Alshamsi, A., Abdallah, S., Rahwan, I., 2009. Multiagent self-organization for a taxi dispatch system, in: 8th International Conference on Autonomous Agents and Multiagent Systems. Citeseer, pp. 21–28.

Anderson, D., 2016. A History of San Francisco's Cab Industry, in Advertisements [WWW Document]. URL http://taxicabsubjects.blogspot.de/2015/05/a-history-of-san-franciscos-cab.html (accessed 1.29.16).

Angelelli, E., Bianchessi, N., Mansini, R., Speranza, M.G., 2009. Comparison of policies in dynamic routing problems. Journal of the Operational Research Society 61, 686–695. doi:10.1057/jors.2009.17

Atasoy, B., Ikeda, T., Ben-Akiva, M.E., 2015. Optimizing a Flexible Mobility on Demand System, in: 94rd Annual Meeting of the Transportation Research Board. pp. 1–21.

Awasthi, A., Chauhan, S.S., Parent, M., Proth, J.M., 2011. Centralized fleet management system for cybernetic transportation. Expert Systems with Applications 38, 3710–3717. doi:10.1016/j.eswa.2010.09.029

Bai, R., Li, J., Atkin, J.A., Kendall, G., 2014. A novel approach to independent taxi scheduling problem based on stable matching. Journal of the Operational Research Society 65, 1501–1510. doi:10.1057/jors.2013.96

Bailey, W.A., Clark, T.D., 1992. Taxi management and route control: a systems study and simulation experiment, in: WSC '92 Proceedings of the 24th Conference on Winter Simulation. pp. 1217–122.

Bailey, W.A., Clark, T.D., 1987. A Simulation Analysis of Demand and Fleet Size effects on taxicab service rates, in: Winter Simulation Conference. pp. 838–844.

Benyahia, I., Potvin, J.-Y., 1997. Genetic programming for vehicle dispatching. IEEE International Conference on Evolutionary Computation 547–552.

Berbeglia, G., Cordeau, J.-F., Laporte, G., 2010. Dynamic pickup and delivery problems. European Journal of Operational Research 202, 8–15. doi:10.1016/j.ejor.2009.04.024

Berger, T., Sallez, Y., Raileanu, S., Tahon, C., Trentesaux, D., Borangiu, T., 2011. Personal Rapid Transit in an open-control framework. Computers and Industrial Engineering 61, 300–312. doi:10.1016/j.cie.2010.12.010

Billhardt, H., Rey, U., Carlos, J., Fernández, A., Rey, U., Carlos, J., Lujak, M., Rey, U., Carlos, J., Ossowski, S., Rey, U., Carlos, J., 2014. Dynamic Coordination in Fleet Management Systems: Toward Smart Cyber Fleets. Intelligent Systems, IEEE 29, 70–76. doi:10.1109/MIS.2014.41

Box, G.E.P., Draper, N.R., 1987. Empirical Model Building and Response Surfaces, in: Empirical Model Building and Response Surfaces. John Wiley & Sons, New York, NY, p. 74 and 424.

Brito, I., Meseguer, P., 2006. Distributed stable marriage with ties and incomplete lists, in: Proceedings of Principles and Practice of Constraint Programming - CP 2006. doi:10.1007/11889205_49

Bullo, F., Frazzoli, E., Pavone, M., 2011. Dynamic Vehicle Routing for Robotic Systems. Proceedings of the IEEE 99.

Cairns, R.D., Liston-Heyes, C., 1996. Competition and regulation in the taxi industry. Journal of Public Economics 59, 1–15. doi:10.1016/0047-2727(94)01495-7

Chang, T.-S., Yen, H.-M., 2012. City-courier routing and scheduling problems. European Journal of Operational Research 223, 489–498. doi:10.1016/j.ejor.2012.06.007

Choo, F.C., Chan, M.C., Ananda, A.L., Peh, L.S., 2012. A Distributed Taxi Advisory System. 2012 12th International Conference on Its Telecommunications (Itst-2012) 193–198. doi:10.1109/ITST.2012.6425165

City and County of San Francisco Taxi Commission Public Convenience & Necessity Report, 2007.

Cordeau, J.-F., Gendreau, M., Laporte, G., Potvin, J.-Y., Semet, F., 2002. A guide to vehicle routing heuristics. Journal of the Operational Research Society 53, 512–522. doi:10.1057/palgrave/jors/2601319

CRAWDAD A Community Resource for Archiving Wireless Data At Dartmouth [WWW Document], 2016. doi:10.15783/C7G88S

D'Orey, P.M., Fernandes, R., Ferreira, M., 2012. Empirical evaluation of a dynamic and distributed taxi-sharing system. 2012 15th International IEEE Conference on Intelligent Transportation Systems 140–146. doi:10.1109/ITSC.2012.6338703

Deflorio, F.P., 2011. Simulation of requests in demand responsive transport systems. IET Intelligent Transport Systems 5, 159. doi:10.1049/iet-its.2010.0026

Dial, R.B., 1995. Autonomous dial-a-ride transit introductory overview. Transportation Research Part C: Emerging Technologies 3, 261–275. doi:10.1016/0968-090X(95)00010-G

Diana, M., 2006. The importance of information flows temporal attributes for the efficient scheduling of dynamic demand responsive transport services. Journal of Advanced Transportation 40, 23–46.

Douglas, G.W., 1972. Price Regulation and Optimal Service Standards: The Taxicab Industry. Journal of Transport Economics and Policy 6, 116–127. doi:10.2307/20052271

Ehmke, J.F., Campbell, A.M., 2014. Customer acceptance mechanisms for home deliveries in metropolitan areas. European Journal of Operational Research 233, 193–207. doi:10.1016/j.ejor.2013.08.028

Eksioglu, B., Vural, A.V., Reisman, A., 2009. The vehicle routing problem: A taxonomic review. Computers & Industrial Engineering 57, 1472–1483. doi:10.1016/j.cie.2009.05.009

Exploratorium Cabspotting [WWW Document], 2016. URL http://www.exploratorium.edu/id/cab.html (accessed 4.19.16).

Fagnant, D., Kockelman, K.M., 2014. The Environmental Implications of Shared Autonomous Vehicles, Using Agent-Based Simulation. Transportation Research Part C: Emerging Technologies 40, 1–18.

Fang, G., Kasim, S., So, S., 2007. San Francisco's taxi dispatch service: Improving reliability and response Report.

Ferrucci, F., Bock, S., 2014. Real-time control of express pickup and delivery processes in a dynamic environment. Transportation Research Part B: Methodological 63, 1–14. doi:10.1016/j.trb.2014.02.001

Ferrucci, F., Bock, S., Gendreau, M., 2013. A pro-active real-time control approach for dynamic vehicle routing problems dealing with the delivery of urgent goods. European Journal of Operational Research 225, 130–141. doi:10.1016/j.ejor.2012.09.016

Firnkorn, J., Müller, M., 2014. Free-floating electric carsharing-fleets in smart cities : The dawning of a post-private car era in urban environments ? Environmental Science & Policy 45, 30–40.

Flores-guri, A.D., Fox, B., Parker, C., Naughten, M., 2012. Local Exclusive Cruising Regulation and Efficiency in Taxicab Markets Local Exclusive Cruising Regulation and Efficiency in Taxicab Markets 39, 155–166.

Frankena, M.W., Pautler, P.A., 1984. An Economic Analysis of Taxicab Regulation. Bureau of Economics, Federal Trade Commission, Washington, D.C.

Furuhata, M., Dessouky, M., Ordóñez, F., Brunet, M., Wang, X., Koenig, S., 2013. Ridesharing :

The state-of-the-art and future directions 57, 28–46.

Gale, D., Shapley, L., 1962. College admissions and the stability of marriage. American Mathematical Monthly 09, 9–15.

Gath, M., Herzog, O., Edelkamp, S., 2014. Autonomous and flexible multiagent systems enhance transport logistics. 2014 11th International Conference and Expo on Emerging Technologies for a Smarter World, CEWIT 2014 0–5. doi:10.1109/CEWIT.2014.7021143

Gendreau, M., Guertin, F., Potvin, J.-Y., Séguin, R., 2006. Neighborhood search heuristics for a dynamic vehicle dispatching problem with pick-ups and deliveries. Transportation Research Part C: Emerging Technologies 14, 157–174. doi:10.1016/j.trc.2006.03.002

Gendreau, M., Guertin, F., Potvin, J.-Y., Taillard, É., 1999. Parallel Tabu Search for Real-Time Vehicle Routing and Dispatching. Transportation Science 33, 381–390. doi:10.1287/trsc.33.4.381

Gendreau, M., Potvin, J.-Y., 1998. Dynamic Vehicle Routing and Dispatching, in: Fleet Management and Logistics. pp. 115–126. doi:10.1007/978-1-4615-5755-5_5

Glaschenko, A., Ivaschenko, A., Rzevski, G., Skobelev, P., 2009. Multi-Agent Real Time Scheduling System for Taxi Companies, in: Proc. of 8th Int. Conf. on Autonomous Agents and Multi-Agent Systems (AAMAS 2009), Budapest, Hungary. pp. 29–36.

Google Books Ngram Viewer [WWW Document], 2013. . Google Books Ngram Viewer. URL https://books.google.com/ngrams (accessed 4.26.16).

He, F., Shen, Z.-J.M., 2015. Modeling taxi services with smartphone-based e-hailing applications. Transportation Research Part C: Emerging Technologies 58, 93–106. doi:10.1016/j.trc.2015.06.023

Horn, M.E.T., 2002. Fleet scheduling and dispatching for demand-responsive passenger services. Transportation Research Part C: Emerging Technologies 10, 35–63. doi:10.1016/S0968-090X(01)00003-1

Ichoua, S., Gendreau, M., Potvin, J.-Y., 2006. Exploiting Knowledge About Future Demands for Real-Time Vehicle Dispatching. Transportation Science 40, 211–225. doi:10.1287/trsc.1050.0114

Kim, H., Oh, J.-S., Jayakrishnan, R., 2005. Effect of Taxi Information System on Efficiency and Quality of Taxi Services. Transportation Research Record: Journal of the Transportation Research Board 1903, 96–104. doi:10.3141/1903-11

Kocer, B., Arslan, A., 2012. Transfer learning in vehicle routing problem for rapid adaptation. International Journal of Innovative Computing, Information and Control 8, 6799–6809.

Kümmel, M., 2014. Real-time Taxi Scheduling Survey [WWW Document]. URL http://goo.gl/forms/NwKwCqx0ay (accessed 4.19.16).

Kümmel, M., Busch, F., Wang, D.Z.W., 2016a. Automated Taxi Vehicle Fleet Scheduling: The Family Model, in: Submitted to 19th Euro Working Group on Transportation Meeting, EWGT 2016.

Kümmel, M., Busch, F., Wang, D.Z.W., 2016b. Taxi dispatching and stable marriage Lessons learned from application on San Francisco taxi data, in: Submitted to hEART2016 5th Symposium Arranged by European Association for Research in Transportation Delft.

Kümmel, M., Busch, F., Wang, D.Z.W., 2015. Taxi Driver's Dilemma : Which Passenger-Selection Strategy Maximizes Profit?, in: Proceedings of Transportation Research Board 94th Annual Meeting 2015.

Kümmel, M., Busch, F., Z.W.Wang, D., 2016c. Booking a taxi in advance – Terra incognita of taxi dispatching. Submitted to European Journal of Operational Research (EJOR).

Kümmel, M., Busch, F., Z.W.Wang, D., 2016d. Taxi dispatching and stable marriage. Procedia Computer Science 83, 163–170. doi:10.1016/j.procs.2016.04.112

Land Transport Authority, 2015. Singapore Land Transport Statistics in Brief 2015 [WWW Document]. URL http://www.lta.gov.sg/content/dam/ltaweb/corp/PublicationsResearch/files/FactsandFigures/Statistics in Brief 2015 FINAL.pdf (accessed 10.8.15).

Land Transport Authority in Singapore, 2014. Land Transport Statistics in Brief. Singapore. doi:10.1007/s11999-008-0346-9

Landratsamt München, 2016. Landkreis München Daten und Fakten [WWW Document]. URL http://www.landkreis-muenchen.de/verwaltung-buergerservice-politik-wahlen/landkreis-muenchen/daten-und-fakten/#c584 (accessed 4.19.16).

Larsen, a, Madsen, O., Solomon, M., 2002. Partially dynamic vehicle routing—models and algorithms. Journal of the Operational Research Society 53, 637–646. doi:10.1057/palgrave.jors.2601352

Larsen, A., Madsen, O.B.G., Solomon, M.M., 2007. Classification Of Dynamic Vehicle Routing Systems, in: Dynamic Fleet Management. doi:10.1007/978-0-387-71722-7_2

Lee, D., Wu, X., Sun, L., 2013. A Limited Information Sharing Strategy for the Taxi-Customer Searching Problem in the Non-Booking Taxi Service, in: Proceedings of the Transportation Research Board 92nd Annual Meeting.

Lee, D.-H., Wang, H., Cheu, R., Teo, S., 2004. A taxi dispatch system based on current demands and real-time traffic conditions. Transportation Research Record 1882, 193–200. doi:10.3141/1882-23

Lee, D.-H., Wu, X., 2013. Dispatching strategies for the taxi-customer searching problem in the booking taxi service, in: Proceedings of the Transportation Research Board 92nd Annual Meeting.

Lee, D.-H., Wu, X., Sun, L., 2013. Limited Information Sharing Strategy for the Taxi-Customer Searching Problem in the Non-Booking Taxi Service. Transportation Research Record: Journal of the Transportation Research Board 2333, 46–54.

Lees-Miller, J.D., Wilson, R.E., 2012. Proactive empty vehicle redistribution for personal rapid transit and taxis. Transportation Planning and Technology 35, 17–30. doi:10.1080/03081060.2012.635414

Liu, L., Andris, C., Ratti, C., 2010. Uncovering cabdrivers' behavior patterns from their digital traces. Computers, Environment and Urban Systems 34, 541–548. doi:10.1016/j.compenvurbsys.2010.07.004

Maciejewski, M., Bischoff, J., 2015. Large-scale microscopic simulation of taxi services. Procedia Computer Science 52, 358–364. doi:10.1016/j.procs.2015.05.107

Maciejewski, M., Bischoff, J., Nagel, K., Berlin, T.U., 2016. An Assignment-Based Approach to Efficient Taxi Dispatching. Intelligent Systems IEEE 31.

McVitie, D., Wilson, L., 1971. The stable marriage problem. Association for Computing Machinery (ACM) 14, 486–490.

Miao, F., Lin, S., Munir, S., Stankovic, J.A., Huang, H., Zhang, D., He, T., Pappas, G.J., 2015. Taxi Dispatch with Real-Time Sensing Data in Metropolitan Areas — a Receding Horizon Control Approach, in: ICCPS '15 Proceedings of the ACM/IEEE Sixth International Conference on Cyber-Physical Systems.

Mitrović-Minić, S., Laporte, G., 2004. Waiting strategies for the dynamic pickup and delivery problem with time windows. Transportation Research Part B: Methodological 38, 635–655. doi:10.1016/j.trb.2003.09.002

Moreira-Matias, L., 2014. On Improving Operational Planning and Control in Public Transportation Networks using Streaming Data: A Machine Learning Approach, in: Proc European Conf. on Machine Learning and Principles and Practice of Knowledge Discovery in Databases - ECML/PKDD. pp. 41–50.

Ngo, M.N., 2004. Fuzzy linear assignment problem: an approach to vehicle fleet deployment, in: Proceedings of the IEEE International Conference on Fuzzy Systems. doi:10.1109/FUZZY.2004.1375583

OpenStreetMap [WWW Document], 2016. . #map=9/37.6882/-122.3245. URL http://www.openstreetmap.org/ (accessed 4.19.16).

Pavone, M., Smith, S.L., Frazzoli, E., Rus, D., 2012. Robotic load balancing for mobility-on-demand systems. The International Journal of Robotics Research 31, 839–854. doi:10.1177/0278364912444766

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

Pillac, V., Guéret, C., Medaglia, A.L., 2011. Dynamic Vehicle Routing Problems: State Of The Art and Prospects. Technical Report 10/4/AUTO <hal-00623474>.

Powell, J., Huang, Y., Bastani, F., Ji, M., 2011. Towards Reducing Taxicab Cruising Time Using Spatio-Temporal Profitability Maps. Proceedings of the 12th International Conference on Advances in Spatial and Temporal Databases 242–260.

Prodhon, C., Prins, C., 2014. A survey of recent research on location-routing problems. European Journal of Operational Research 238, 1–17. doi:10.1016/j.ejor.2014.01.005

Psaraftis, H.N., 1980. A Dynamic Programming Solution to the Single Vehicle Many-to-Many Immediate Request Dial-a-Ride Problem. Transportation Science. doi:10.1287/trsc.14.2.130

Ritzinger, U., Puchinger, J., Hartl, R.F., 2015. A survey on dynamic and stochastic vehicle routing problems. International Journal of Production Research 7543, 1–17. doi:10.1080/00207543.2015.1043403

Ronald, N., Thompson, R., Winter, S., 2015. Simulating Demand-responsive Transportation: A Review of Agent-based Approaches. Transport Reviews 35, 404–421. doi:10.1080/01441647.2015.1017749

Roth, A.E., Rothblum, U.G., Vande Vate, J.H., 1993. Stable Matchings, Optimal Assignments, and Linear Programming. Mathematics of Operations Research.

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

Santi, P., Resta, G., Szell, M., Sobolevsky, S., Strogatz, S.H., Ratti, C., 2014. Quantifying the benefits of vehicle pooling with shareability networks. Proceedings of the National Academy of Sciences 111, 13290–13294. doi:10.1073/pnas.1403657111

Seow, K.T., Dang, N.H., Lee, D.-H., 2010. A Collaborative Multiagent Taxi-Dispatch System. IEEE Transactions on Automation Science and Engineering 7, 607–616. doi:10.1109/TASE.2009.2028577

Seow, K.T., Lee, D.H., 2010. Performance of multiagent taxi dispatch on extended-runtime taxi availability: A simulation study. IEEE Transactions on Intelligent Transportation Systems 11, 231–236. doi:10.1109/TITS.2009.2033128

Shamma, J., 2008. Task assignment for mobile agents, in: Cooperative Control of Multi-Agent Systems. pp. 109–137.

Shi, W., Tong, C.O., Wong, S.C., 2010. Influence of Real-Time Information Provision to Vacant Taxi Driver on Taxi System Performance. Proceeding of the Twelfth World Conference of Transport Research 1–25.

Shrivastava, M., Chande, P.K., Monga, A.S., Kashiwagi, H., 1997. Taxi despatch: a fuzzy rule approach, in: Proceedings of Conference on Intelligent Transportation Systems. pp. 978–982. doi:10.1109/itsc.1997.660606

Silva, G.R., Mateus, A.P., 2003. A mobile location-based vehicle fleet management service application. IEEE IV2003 Intelligent Vehicles Symposium 25–30. doi:10.1109/IVS.2003.1212877

Sirisoma, R.M.N.T., Lam, W.H.K., Yang, H., Wong, S.C., Wang, D., Zhang, P., 2010. Empirical Evidence for Taxi Customer-Search Model. Proceedings of the Institution of Civil Engineers: Transport 163, 203–210. doi:10.1680/tran.2010.163.4.203

Solution Methods for VRP [WWW Document], 2013. . Neo research group, University of Malaga. URL http://neo.lcc.uma.es/vrp/solution-methods/

Teo, C.-P., Sethuraman, J., Tan, W.-P., 2001. Gale-Shapley Stable Marriage Problem Revisited: Strategic Issues and Applications. Management Science 47, 1252–1267. doi:10.1287/mnsc.47.9.1252.9784

Thomopoulos, S.C.A., Argyreas, N., Gorgas, V., Kyriazanos, D., Soulos, G., Makri, E., 2007. ToD

Pilot: A Location-based Taxi on Demand System Demonstrator, in: 2007 IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications. Ieee, pp. 1–7. doi:10.1109/PIMRC.2007.4394876

Treleaven, K., Pavone, M., Frazzoli, E., 2012. Asymptotically Optimal Algorithms for Pickup and Delivery Problems with Application to {Large-Scale} Transportation Systems (submitted). arXiv:1202.1327 1–36. doi:10.1109/TAC.2013.2259993

van Hentenryck, P., Bent, R., Upfal, E., 2010. Online stochastic optimization under time constraints, Annals of Operations Research. doi:10.1007/s10479-009-0605-5

Wang, F., Yang, M., Yang, R., 2008. Simulation of Multi-Agent based Cybernetic Transportation System. Simulation Modelling Practice and Theory 16, 1606–1614. doi:10.1016/j.simpat.2008.08.008

Wang, H., Lee, D.-H., Cheu, R., 2009. PDPTW Based Taxi Dispatch Modeling for Booking Service. 2009 Fifth International Conference on Natural Computation 242–247. doi:10.1109/ICNC.2009.676

Wilson, N.H.M., Sussman, J.M., Goodman, L.A., Hignnet, B.T., 1969. Simulation of a Computer Aided Routing System (CARS), in: Proceedings of the Third Conference on Applications of Simulation. pp. 171–183. doi:10.1007/s13398-014-0173-7.2

Wu, X., Lee, D., 2013. An Integrated Taxi Dispatching Strategy Handling both Current and Advance Bookings. Journal of the Eastern Asia Society for Transportation Studies 10, 1836–1855.

Xin, C.-L., Ma, W.-M., 2004. Scheduling for on-line taxi problem on a real line and competitive algorithms. Proceedings of 2004 International Conference on Machine Learning and Cybernetics (IEEE Cat. No.04EX826) 5, 26–29. doi:10.1109/ICMLC.2004.1378561

Yang, H., 2010. Equilibrium Properties of Taxi Market with Search Frictions. Transport Economics.

Yang, H., Leung, C., Wong, S., Bell, M., 2010. Equilibria of bilateral taxi–customer searching and meeting on networks. Transportation Research Part B: Methodological 44, 1067–1083. doi:10.1016/j.trb.2009.12.010

Yang, H., Yang, T., 2011. Equilibrium properties of taxi markets with search frictions. Transportation Research Part B: Methodological 45, 696–713. doi:10.1016/j.trb.2011.01.002

Yang, T., Yang, H., Wong, S.C., Bay, C.W., Kong, H., 2010. Modeling Taxi Services with a Bilateral Taxi-Customer Searching and Meeting Function. Science And Technology.

Yuan, J., Zheng, Y., Zhang, L., Xie, X., Sun, G., 2011. Where to Find my Next Passenger?, in: Proceedings of the 13th ACM International Conference on Ubiquitous Computing. pp. 1–10.

Yuan, N.J., Zheng, Y., Zhang, L., Xie, X., 2013. T-Finder : A Recommender System for Finding Passengers and Vacant Taxis. IEEE Transactions on Knowledge & Data Engineering 25, 2390–2403. doi:10.1109/TKDE.2012.153

Zachariah, J., Gao, J., Kornhauser, A., Mufti, T., 2014. Uncongested Mobility for All: A Proposal for an Area Wide Autonomous Taxi System in New Jersey, in: 93rd Annual Meeting of the Transportation Research Board. pp. 1–14.

Zhan, X., Qian, X., Ukkusuri, S. V, 2014. A Graph Based Approach to Measuring the Efficiency of Urban Taxi Service System, in: Transportation Research Board Conference. pp. 1–13.

Zhang, D., He, T., 2012. pCruise: Reducing Cruising Miles for Taxicab Networks, in: 2012 IEEE 33rd Real-Time Systems Symposium. Ieee, pp. 85–94. doi:10.1109/RTSS.2012.61

Glossary and abbreviations

DOD (Degree of Dynamism) – Ratio of immediate booking requests from all booking requests. If DOD=100% all booking requests are immediate. If DOD=0% all booking requests are advance.

FCFS (First-come, first-served) – Assignment strategy that assigns passenger booking requests sequentially

GPS (Global Positioning System) – A space-based navigation system that provides location information on the Earth.

ID (An identifier) – An unique identifier used to describe unique passenger requests and unique taxis

KPI (Key performance indicator) – A type of performance measurement such as achievement of some levels of operational goals.

Log – An outcome of a logger, which saves the position of GPS device (and other variables) at (regular) time intervals.

OSM (OpenStreetMap) – A collaborative project to create a free editable map of the world.

MATLAB (matrix laboratory) – A multi-paradigm numerical computing environment and fourth-generation programming language developed by MathWorks.

SM (Stable marriage) – A stable marriage problem (also stable matching problem) is the problem of finding a stable matching between two sets of elements given an ordering of preferences for each element.

Unix time – A system for describing instants in time, defined as the number of seconds that have elapsed since 00:00:00 Coordinated Universal Time, Thursday, 1 January 1970.

*VRP (Vehicle Routing Problem) – A combinatorial optimization and integer programming problem, which asks "What is the optimal set of routes for a fleet of vehicles to traverse in order to deliver to a given set of customers?" See chapter 2.1 ”*The taxi dispatching problem*” for more detail.*

Terms used in almost an interchangeable manner in taxi research literature and this thesis:

- Approach / experiment
- Assignment / insertion
- Booking request / booking / passenger / customer / client / trip / demand / pax
- Destination / drop-off positions
- Dispatching / scheduling / management / matching
- Dispatching algorithm / dispatching strategy
- Length of decision epoch / buffering time interval length / matching time
- Location / position
- Occupied / in service / on shift / on-trip
- Origin / pick-up positions
- Re-assignment / re-insertion / re-scheduling / re-optimization
- Sequential / one-by-one / FCFS
- Simultaneous assignment / concurrent assignment / group assignment
- Taxis / drivers / supply
- Taxi schedule / taxi plan
- Trip travel time / on-trip time
- Vacant / empty / idle

List of figures

List of tables

Summary of publications

Core publications and preprints related to this thesis:

Taxi dispatching and stable marriage
Michal Kümmel, Fritz Busch, David Z.W. Wang
Procedia Computer Science Journal 2016 (DOI 10.1016/j.procs.2016.04.112)

Taxi dispatching and stable marriage
Michal Kümmel, Fritz Busch, David Z.W. Wang
The 7th International Conference on Ambient Systems, Networks and Technologies (ANT 2016)

Publications and preprints extending this thesis:

Framework for automated taxi operation: The family model
Michal Kümmel, Fritz Busch, David Z.W. Wang
In Proceedings of 19th Euro Working Group on Transportation Meeting, EWGT 2016

Taxi Driver's Dilemma: Which Passenger-Selection Strategy Maximizes Profit?
Michal Kuemmel, Fritz Busch, David Z. W. Wang
In Proceedings of Transportation Research Board 94th Annual Meeting 2015

Not yet published publications / publications in writing:

Taxi dispatching and stable marriage Lessons learned from application on San Francisco taxi data
Michal Kuemmel, Fritz Busch, David Z. W. Wang

Booking a taxi in advance - Terra incognita of taxi dispatching
Michal Kuemmel, Fritz Busch, David Z. W. Wang

Other publications not included in this thesis:

Supply-demand optimization for electric taxis in Singapore
M. Kümmel, F. Busch, A. Rau, Y. D. Wong
In Proceedings of 23. Verkehrswissenschaftliche Tage 2012