Designing taxi ridesharing systems with shared pick-up and drop-off locations: Insights from a computational study

Miriam Stumpe[a,*], Peter Dieter[a], Guido Schryen[a], Oliver Müller[a], Daniel Beverungen[a]

[a] *Paderborn University, Warburger Str. 100, 33098 Paderborn, Germany*

Abstract

Taxi ridesharing (TRS) systems are considered one means towards more sustainable transportation by increasing car occupancy rates and thereby significantly improving the efficiency of urban transportation systems. In this study, we consider TRS with shared pick-up and drop-off locations, where customers of a shared trip might be required to walk a short distance from their origin/to their destination. Related research has discussed the advantages of this approach over other TRS variants, including shorter travel times, lower fuel consumption and fewer emissions. However, these studies do not investigate how a TRS ought to be designed under different environmental conditions to maximize its effectiveness in terms of rejection and sharing rate. We contribute to closing this gap with three achievements. First, we propose a new mathematical model that provides a conceptualization of the TRS problem with shared pick-up and drop-off locations. Second, we implement a rolling horizon approach and conduct extensive computational experiments based on empirical data from New York City and Porto. In our experiments, we vary and combine several exogeneous (environmental) and design-oriented factors and show that both exert considerable influence on the rejection rate, sharing rate and service quality.

*Corresponding author.

Email addresses: `miriam.stumpe@upb.de` (Miriam Stumpe), `peter.dieter@upb.de` (Peter Dieter), `guido.schryen@upb.de` (Guido Schryen), `oliver.mueller@upb.de` (Oliver Müller), `daniel.beverungen@upb.de` (Daniel Beverungen)

Third, for practitioners considering a TRS with shared pick-up and drop-off locations, our guidelines highlight the importance of system design, particularly in leveraging extended waiting times to attain low rejection rates and foster high sharing rates.
Keywords: taxi ridesharing, sustainable transport, system design

1. Introduction

For a long time, people traveling in urban areas had to choose between private transport, including taxis and private cars, and public transport, including buses, trams and subways. The resulting dichotomy decision options have been equated to a trade-off between cost and comfort: While private transportation came with high comfort at high cost, public transportation were involved with low comfort at low cost. While this dichotomy still exists, being the two dominant travel modes in most regions, nowadays, a third travel mode paradigm appeared with the rise of ridesharing services such as Uber, Lyft, and DiDi, whose on-demand services are based upon matching individual drivers and their cars with customers who need transport (Weber, 2020; Silwal et al., 2019; Furuhata et al., 2013). Taxi ridesharing (TRS) systems (which can be seen as *shared-ride mobility-on-demand systems*) are a response to the resulting market pressure on taxi companies, offering the flexibility and speed of private cars but at lower costs, due to the higher utilization of vehicles and drivers. This mode of transportation is characterized by matching several separate taxi trip requests with similar spatio-temporal characteristics into a shared taxi trip (Barann et al., 2017). Along with increased service options for customers, TRS systems are considered a promising means towards more sustainable and more efficient urban transportation by reducing fuel consumption and air pollution, and increasing car occupancy rates and reducing traffic congestion, respectively (Agatz et al., 2012; Santi et al., 2014).

The landscape of TRS systems is pervaded by a variety of systems which differ in several regards. A frequently used classification distinguishes different numbers of pick-up locations (one or many) and the number of drop-off locations (one or many) per taxi trip, resulting in four classes of TRS approaches (Ting et al., 2021; Smet, 2021; Bathla et al., 2018). In its traditional configuration, the *many-to-many* approach corresponds to a door-to-door transport, where customers are picked up and dropped off at exactly their desired departure and arrival locations. Such types of systems do not only require matching requests of customers who may share a joint trip but also determining orders in which customers are picked up and dropped off (routing). While customers benefit from being picked up and dropped off at the desired locations, they may suffer from detours, which, in turn, reduce their convenience. With the *one-to-one* TRS approach, matching customer requests needs to be complemented by determining common or nearby pick-up and drop-off locations for shared trips. Here, customers need to accept short walking distances before and after a shared taxi trip but do not have to tolerate detours of the taxi. More direct vehicle routes, enabled by short walking distances, increase system efficiency by reducing vehicle distance traveled, energy consumption, emissions, and operating costs, by further leveraging the benefits of sharing rides compared to the traditional many-to-many approach (Stiglic et al., 2016). Furthermore, customers enjoy higher convenience through shorter travel times by detour avoidance (Stiglic et al., 2015).

A real-world example of a product based on the above considerations is Uber Express Pool, which offers customers lower fares based on lower operating costs (Stock, 2018). Scenarios in which one-to-one TRS may be favorable include situations in which guests of a single hotel or several hotels closely located to each other have similar destinations in the city center (e.g., a shopping mall, a train station or a sightseeing spot), or in which participants of an event (e.g., a farmer's market or a party) want to return home to the same neighborhood.

While the many-to-many TRS approach has received much attention in the literature (Mourad et al., 2019; Liang et al., 2020; Wang and Li, 2021; Ting et al., 2021) considerably less attention has been paid to one-to-one TRS systems. Works dedicated to these systems include the studies of Barann et al. (2017) and Qian et al. (2017), which identify high sharing potentials in the considered example of New York City (NYC) for a deterministic setting where all requests of a planning period are known in advance. Moreover, several studies Wang et al. (2019); Yan et al. (2020) demonstrate that one-to-one systems might be superior to many-to-many systems in terms of, for example, shorter waiting and service times for customers, less energy consumption and reduced kilometers traveled by the vehicle fleet, leading to a more sustainable system.

However, the literature lacks studies that consider a one-to-one TRS in a dynamic setting where information reveals itself during the planning period. In addition, the literature is silent on insights into how system design factors of one-to-one TRS systems, such as the maximum walking distance allowed, and exogenous (environmental) factors, such as the temporal-spatial density of customer requests, determine the effectiveness of the TRS system (e.g., in terms of rejection and sharing rate). Since the general concept of one-to-one TRS has been proven to be beneficial, gaining insights into the impact of both exogenous and system design factors on the effectiveness of one-to-one TRS is valuable for providers of TRS systems in terms of design decisions.

Addressing the overall research question of how TRS systems with shared pick-up and drop-off locations should be designed, the contributions of our work are threefold: 1) We propose a mathematical formulation of the one-to-one TRS problem in the form of an integer linear program (ILP). The objective is to serve as many requests as possible while minimizing the overall cost for vehicle-trip assignments, considering the perspectives of both customers and operators, all within the constraints of a limited taxi fleet. Problem dynamics regarding future customer requests are accounted for by using a rolling horizon solution approach. 2) We conduct extensive computational experiments based upon real-world data from NYC and Porto. We also perform a sensitivity analysis to investigate the effects of various system design factors and exogenous parameter values on the rejection rate, sharing rate and service quality. 3) Based upon our computational insights, we derive guidelines for practitioners who plan to design and implement one-to-one TRS systems.

The remainder of this paper is organized as follows. Section 2 provides selected prior works on one-to-one TRS that are related to our work. In Section 3, we describe the problem and provide an illustrative example. In Section 4, we develop the mathematical model and present the solution approach. In Section 5, we describe our computational study by presenting the experimental design and computational results. We discuss the results and derive managerial guidelines in Section 6. Finally, we conclude with a summary of our study and an outlook on future research in Section 7.

2. Related work

Shared-ride services, i.e. different concepts of people sharing rides, have been intensively discussed in the operations research community; for literature reviews, we refer to Agatz et al. (2012), Mourad et al. (2019) and Tafreshian et al. (2020). In our work, we investigate the concept of TRS with shared pick-up an drop-off locations (meeting points). Despite the limited literature on this specific combination, there is a substantial body of knowledge on related problems. In the following, we briefly sketch the literature landscape of TRS and, then, give an overview of literature on shared-ride services with meeting points.

2.1. Taxi ridesharing

TRS systems aim at bringing together customers with similar spatio-temporal transportation needs, satisfying their on-demand customer requests through shared trips executed by a centrally organized vehicle fleet (Silwal et al., 2019). TRS is related to two other shared-ride services, from which it is distinguished as follows. Focusing on single, non-recurring trips, TRS differs from *carpooling*, which requires a long-term commitment between two or more people (Correia and Viegas, 2009; Bruck et al., 2017). In contrast, the dynamic problem setting in TRS is also a characteristic of most *peer-to-peer (P2P)* ridesharing systems (Sun et al., 2020). However, in P2P ridesharing, private drivers dynamically enter a platform to share their trips with riders who have similar itineraries, while a TRS system has access to a centrally organized taxi fleet and, therefore, tends to be better informed about where and when individual taxis become available. In addition, P2P ridesharing often focuses on bilateral matching between one driver and one rider, while TRS always involves multiple customers sharing the same ride, also referred to as ridepooling or ridesplitting.

Operational issues of TRS systems are two-fold: First, idle taxis need to be prepositioned or rebalanced towards regions with a high rate of unserved customers (Yuan et al., 2011; Li et al., 2018a; Demissie et al., 2020). Second, customers need to be grouped and assigned to taxis for which routes need to be determined. In this work, we focus on the grouping of customer requests and continue presenting related approaches suggested in literature.

The TRS problem originated as a dynamic Dial-a-ride problem (DARP) with multiple vehicles (Ma et al., 2013; Hosni et al., 2014) that consists of designing minimum cost vehicle routes and schedules for customers who specify pickup and delivery requests between origins and destinations (Cordeau and Laporte, 2007). The dynamic problem setting, i.e., requests arrive dynamically over time and are not known in advance, is one key characteristic that resulted in several heuristic solution methods for centralized (Ma et al., 2013; Hosni et al., 2014; Santos and Xavier, 2015; Jung et al., 2016) and decentralized TRS problems (d'Orey et al., 2012; Bathla et al., 2018; Manjunath et al., 2021). Due to TRS's high complexity, mainly caused by the routing part of the problem, only a few exact solution methods have been proposed. Santi et al. (2014), Alonso-Mora et al. (2017) and Liang et al. (2020) develop methods based on a rolling horizon approach where an optimization model is used to periodically determine optimal taxi assignments and routes for all collected requests during a specified time window. Wang et al. (2018) and Kucharski and Cats (2020) also consider a rolling horizon approach, albeit for P2P ridesharing. Their methodologies focus on matching customer requests to shared trips, wherein routing is implicitly handled by identifying optimal matches between customer requests and shared trips with predefined routes. In our work, we adopt a similar rolling horizon approach, by matching passenger requests to taxi trips, implicitly optimizing taxi routes within a dynamic one-to-one TRS system.

2.2. Shared-ride services with meeting points

While all of the above approaches consider door-to-door transportation, another literature stream examines shared-ride services where customers must walk short distances to common meeting points to be picked up for transport. For a more extensive review on optimizing shared-ride services with walking legs, we refer to Wang et al. (2022). Evidence shows that the introduction of meeting points in shared-ride services has multiple advantages: 1) Decreasing extra detours for customers and, thus, reducing inconvenience (Stiglic et al., 2015; Zhao et al., 2018; Li et al., 2018b), 2) increasing the sharing potential and, therefore, improving the overall efficiency of the system and mitigating emissions and environmental impact (Stiglic et al., 2015; Smet, 2021; Fielbaum et al., 2021), and 3) offering more privacy to customers, as their origin and destination remain unknown to other customers (Aïvodji et al., 2016; Khazbak et al., 2020).

However, the literature on meeting points in taxi ridesharing systems is rather limited and mostly neglects the dynamic nature of the problem. Therefore, in this section, we also consider the literature on deterministic problem settings and other shared-ride services with meeting points, such as carpooling or P2P ridesharing. Another shared-ride service that is often studied in the context of meeting points is long-distance ridesharing (interurban) (Czioska et al., 2017; Chen et al., 2019), which we do not consider, since this problem does not share the main characteristics of taxi ridesharing, such as requests arriving at short notice.

Barann et al. (2017) and Qian et al. (2017) assume a deterministic setting where all necessary data is known at the moment of planning. Methods of both approaches determine the meeting points on-the-fly during optimization depending on the requested departure locations of customers. Barann et al. (2017) apply an event-driven greedy heuristic that searches for matching candidates for each incoming request, demonstrating savings of up to 500 tons of carbon emissions per week in NYC. Qian et al. (2017) present an optimization model and two solution algorithms (exact and heuristic) for grouping similar customer requests into shared taxi trips. The authors compare the performance of both algorithms and explore pricing policies that provide incentives for customers to share, showing that this can save up to 47% of total travel miles if applied to NYC taxi trip data. These two approaches are referred to as *one-to-*

one TRS, since shared taxi trips have one common pick-up and drop-off location for onboarding and unboarding of all customers without any intermediate stops. Aissat and Oulamara (2014) and Yan et al. (2020) examine the one-to-one approach for P2P ridesharing. Aissat and Oulamara (2014) propose exact and heuristic approaches to identify the pick-up and drop-off locations that minimize the total travel costs of a driver and a rider. Yan et al. (2020) demonstrate the benefits of the one-to-one approach by performing an equilibrium analysis, which shows an improved rider and driver experience in case of demand spikes.

Other approaches consider shared-ride services with predefined meeting points allowing for intermediate stops to pick up or drop off customers, thus referred to as *extended many-to-many.* By restricting the number of pick-up and drop-off locations, the extended many-to-many TRS can be simplified into a one-to-one TRS. Therefore, the extended many-to-many TRS can be regarded as a generalized form, incorporating the one-to-one TRS as a special case. This does not apply to the traditional many-to-many TRS, as this and the one-to-one TRS are two distinct problems. Ham (2021) and Aliari and Haghani (2022) suggest optimization models to find optimal taxi routes while selecting the best pick-up and drop-off location for each customer, assuming a deterministic problem setting. Ham (2021) proposes a constraint programming approach and a two-phase approach based on a warm start to improve performance for solving the optimization model. The model of Aliari and Haghani (2022) solves small problem instances previously formed by a pre-matching heuristic for clustering similar requests. Further clustering algorithms are presented by Martínez et al. (2015) and Czioska et al. (2019), which form small groups of customers for a minibus service, stopping at one meeting point per cluster. Balardino and Santos (2016) and Miklas-Kalczynska and Kalczynski (2021) consider carpooling with meeting points, where each participant has a specific point of origin and the final destination is common for all participants, which corresponds to a many-to-one approach according to our categorization. Miklas-Kalczynska and Kalczynski (2021) show that adding the option of meeting points improves system-wide driving distance savings.

While all abovementioned studies are deterministic, only few papers consider a dynamic setting with stochastically incoming requests for shared-ride services with meeting points. There are some event-driven approaches that perform a planning operation each time a new customer request is received. For example, the TRS system considered by Lyu et al. (2019) generates for each request a sharing schedule consisting of a set of companions, the shortest sharing route and the best pick-up/drop-off locations by maximizing the satisfaction of involved customers. Gökay et al. (2019) map the pick-up and drop-off locations of a new request to those of accepted requests in a vehicle schedule and extend the problem to include a service provider's decision to reject or accept a request. Engelhardt and Bogenberger (2021) propose a simulation framework designed to investigate the influence of boarding locations on TRS systems. Fielbaum et al. (2021) build upon the rolling horizon method proposed by Alonso-Mora et al. (2017) and adapt it to include meeting points. Similar to other studies (Li et al., 2019; Lyu et al., 2019), the authors compare the shared-ride service with meeting points to a classical door-to-door transport and demonstrate additional travel distance savings through the introduction of meeting points. However, none of the studies on dynamic shared-ride systems with meeting points consider a one-to-one approach.

To the best of our knowledge, none of the presented studies consider one-to-one TRS systems in a dynamic setting, although the dynamic nature is one main characteristics of taxi ridesharing. Moreover, there is a lack of studies on how to design TRS systems (e.g., setting the optimization frequency) in dependence on several exogenous factors, such as the density of requests. Drawing on both observations, we contribute to the literature by investigating the TRS problem with shared pick-up and drop-off locations where customer requests are not previously known but revealed over time.

3. Problem description

We consider the main operations problem of a one-to-one TRS system. It optimizes grouping multiple customer requests with similar spatio-temporal characteristics into a shared taxi trip, assuming customers' willingness to walk short distances to or from shared pick-up and drop-off locations. The objective is to serve as many requests as possible while minimizing the waiting and walking times of customers and the driving times of taxis, all within the constraints of a limited taxi fleet.

A customer request refers to a transportation need between two locations of one or more passengers at a given time. We consider the case where such requests dynamically enter a central system operated by a taxi provider owning a limited vehicle fleet and employing dedicated drivers. The taxi provider has to decide whether to service a request as a traditional or a shared taxi trip and assign requests to taxi vehicles. If no vehicle can reach the desired departure location of the request in time, it has to be rejected. Since a shared taxi trip is characterized by only one departure location and one arrival location but serves multiple requests simultaneously, customers might need to walk short distances from/to their requested locations. Intermediate stops during a trip are not considered as they might lead to delays for customers due to detours. A shared taxi trip has to ensure a promised level of service for all included requests: Customers receive information about their taxi trips within reasonable time before the desired departure time. Thus, the taxi provider can use the time between the requests' entry and the notification to the customer for processing requests. The available time for processing a request can vary significantly depending on the customers' preferences about the minimum notification time. The system addresses medium and short-term requests, e.g., customers can announce their requests a few hours to a few minutes before their desired departure time.

We assume a limited, homogeneous taxi fleet with $|V|$ being the number of available vehicles in the fleet. It can happen that for some requests no taxi is available on time, so that they need to be rejected. Regarding customer behavior, we make simplifying assumptions such as comprehensive sharing-willingness. We further suppose that all customers arrive at the specified departure locations on time, walking there with the same average walking speed. We note that this assumption is unlikely to hold in reality and late customer arrivals can degrade the system. When establishing shared taxi trips we do not address customers' particular preferences, such as preferred gender, age, or habits (e.g., non-smoking) of fellow customers.

Based on the above assumptions, the task is to determine taxi trips (in terms of departure time and locations for departure and arrival) and assign them to vehicles of a limited taxi fleet, maximizing served requests and minimizing the total cost for customers and operators, subject to several trip-sharing constraints.

Figure 1 illustrates a problem instance with four customer requests and a sample solution with two taxi trips, where triangles and squares represent the departure and arrival data per request, respectively. The spatial perspective represents the departure and arrival locations of the respective requests in a grid. Here, a segment length corresponds to the maximal allowed walking distance of a customer from the requested departure location to the origin of a taxi trip and from its destination to the requested arrival location. The temporal perspective shows a timeline per request, including the requested (top) and realized (bottom) departure and arrival times. The depicted solution includes two trips consisting of one traditional and one shared taxi trip. As the requested departure of the request marked in white is further than one segment length away from all other departures, a traditional taxi trip serves this request. The remaining three requests fulfill all trip-sharing constraints resulting in a shared taxi trip which requires two requests to walk short distances to/from the joint departure and arrival locations of the shared trip. By taking into account the duration for walking to the departure location, the shared trip starts at 11:09, a few minutes after the desired departure times of all requests involved. As a result, in this example, trip-sharing saves two taxis and around half of the driving distance consumed by a traditional taxi system.

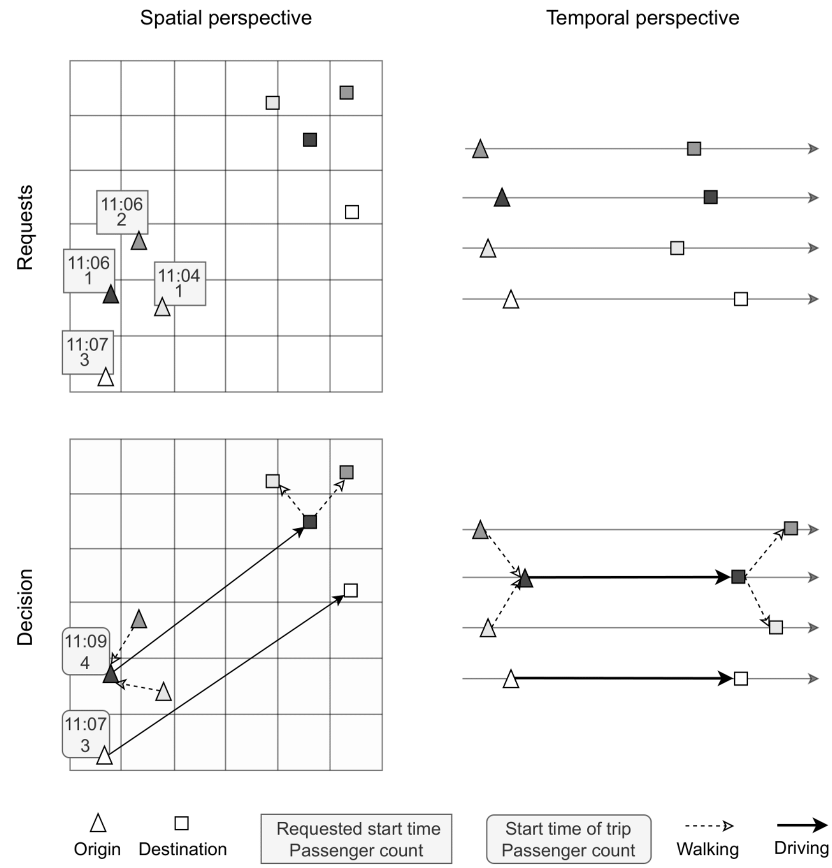

Figure 1: Illustrative Example

Furthermore, to account for the dynamic nature of the problem, i.e., requests arrive sequentially over time, we propose a system that uses a rolling horizon approach to find optimal taxi trips. As argued in Section 2, this is a common approach in literature on shared-ride services (Tafreshian et al., 2020), where it is also referred to as ‘batch approach’ (Kucharski and Cats, 2020) or ‘rolling horizon strategy’ (Agatz et al., 2011). Therefore, we split an entire planning period (such as a day) into many deterministic batches, further referred to as *optimization batches*.

4. Mathematical model and solution framework

In this section, we first introduce the reader to the problem specifications by explaining the notions of *requests* and *taxi trips*. We then describe the constraints needed to create a set of feasible taxi trips including the vehicle assignment. Furthermore, we formulate the mathematical model as an integer linear program (ILP). Finally, we suggest a rolling horizon framework as the overall solution approach. This approach involves solving a sequence of instances of the model iteratively to account for the dynamic problem nature by decomposing the overall planning horizon into planning batches of shorter duration.

4.1. Notions

Request. A request i is characterized by the associated number of passengers $\hat{p}_i$[1], the requested departure and arrival location ($\hat{l}_i^{dep}$ and $\hat{l}_i^{arr}$), the costs $\hat{c}_i$ incurred for a traditional taxi trip serving request i, the requested departure time $\hat{t}_i^{dep}$ and the time when it enters the system, denoted by request time $\hat{t}_i^{req}$. The costs are measured in USD and are usually composed of a base fare and a variable fare, which depends on the distance traveled, the time of the day, and other surcharges (NYC City Taxi & Limousine Commission: TLC, 2023). In addition, we introduce additional temporal concepts, such as the notification time $\hat{t}_i^{not}$, which is the latest possible time the customer would like to be informed about his/her taxi trip. From the departure time and the geographic information, we can derive an expected arrival time $\hat{t}_i^{arr}$ assuming normal traffic conditions. The lead time λ_i of a request expresses its spontaneity, depicting the difference between the requested departure time and the request time. The time flexibility of a request consists of three parts. First, the minimum notification time α, i.e., the minimum duration between a customer getting a notification and wishing to start the trip. Second, the maximum waiting time β represents the difference between the latest accepted departure time of the taxi and the requested departure time. This time is available for a walking distance that

[1] A hat denotes the notation of a request to distinguish it from the notation of a taxi trip.

may be necessary. Third, the maximum delay γ outlines the latest accepted arrival time. Figure 2 provides an overview of the relation between the presented temporal concepts of a request i.

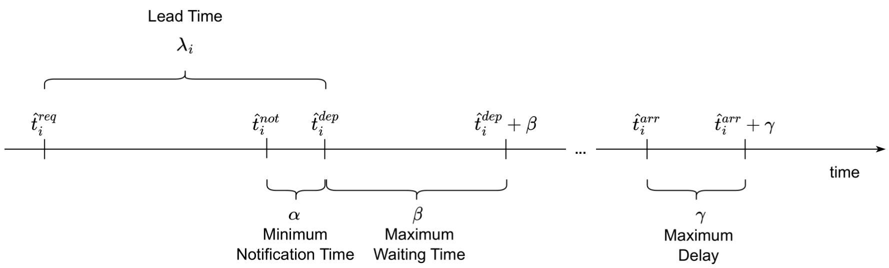

Figure 2: Temporal parameters of a request i

Taxi trip. A taxi trip k represents the realization of one or more requests i by the travel of taxi vehicle v. Thus, for each taxi trip, there is a set $\mathcal{I}_k$, which contains all requests served by trip k. This set specifies whether the trip is a traditional ($|\mathcal{I}_k| = 1$) or a shared ($|\mathcal{I}_k| > 1$) taxi trip. Similar to a request, l_k^{dep} and l_k^{arr} stand for the departure and arrival location of taxi trip k, respectively, while t_k^{dep} and t_k^{arr} display departure and arrival time. The number of passengers is given by p_k. The parameter c_k specifies the cost per request for taxi trip k, assuring that the number of all contained requests times this cost ($|\mathcal{I}_k| \cdot c_k$) corresponds to the earnings of the taxi provider for the taxi trip. Most important for the optimization of taxi trips is parameter $cost(v, k)$ as it represents the costs associated with the assignment of taxi vehicle v to (shared) trip k (see Equation (1) where θ_w and θ_d are functions which measure the walking and driving duration between two locations, respectively).

$$
\begin{aligned}
cost(v,k) = \sum_{i \in \mathcal{I}_k} \hat{p}_i \Big[& u_b(t_k^{dep} - \hat{t}_i^{dep}) + u_w \theta_w(\hat{l}_i^{dep}, l_k^{dep}) + u_d \theta_d(l_k^{dep}, l_k^{arr}) \\
& - \theta_d(\hat{l}_i^{dep}, \hat{l}_i^{arr}) \Big] + u_o \Big[\theta_d(l_{v,k}^{taxi}, l_k^{dep}) + \theta_d(l_k^{dep}, l_k^{arr}) \Big] \qquad (1)
\end{aligned}
$$

Similar to Fielbaum et al. (2021), we define the costs for a vehicle-trip-assignment as the sum of customers' costs and operators' costs. The former encompasses waiting, walking, and in-vehicle travel costs, while the latter arises from the distance driven by the vehicles. Consequently, u_b, u_w and u_d are the unitary values of waiting, walking and traveling over the vehicle, respectively, and u_o is the unitary cost of operating a vehicle. For each request contained in trip k, we subtract the driving duration of a conventional taxi trip between the requested origin and destination $\theta_d(\hat{l}_i^{dep}, \hat{l}_i^{arr})$ from the in-vehicle time to reward time savings achieved through ride-sharing.

4.2. Creation of feasible trips

To create a set of all feasible taxi trips K, all of the following constraints must be met. These constraints ensure a certain level of customer service. We note that the number of trips is subject to factorial growth. However, the assumptions which will be described in the following allow us to filter trips efficiently and consequently, drastically reduce the number and time to create all feasible trips even for large realistic instances. For reasons of complexity and consequently, computational efficiency, pick-up and drop-off locations of a shared trip always correspond to desired departure and arrival locations of at least one of the requests included in the shared trip.

Occupancy constraints. Two constraints limit the occupancy of a taxi and we assume that a request i can include multiple passengers ($\hat{p}_i > 1$). First, the parameter r^{max} limits the maximum number of customer requests that can be merged into a shared trip k (Constraint 2). This sets a limit on the number of foreign parties per trip, represented by $|\mathcal{I}_k|$. Second, Constraint 3 prohibits the number of passengers of a taxi trip k from exceeding the seating capacity q of a taxi. Given a customer request i including multiple passengers, $|\mathcal{I}_k| \neq p_k$ holds for all feasible taxi trips serving this request.

$$|\mathcal{I}_k| \leq r^{max} \qquad \forall\, k \in K \tag{2}$$

$$p_k \leq q \qquad \forall\, k \in K \tag{3}$$

Temporal constraints. A shared trip needs to fulfill three temporal constraints to be feasible. The departure time of trip k (Constraint 4) and the arrival time of taxi v (Constraint 5) must be within the maximum waiting time β of all served requests. Moreover, Constraint 6 limits the walking time of each request from the requested departure location to the joint departure of trip k by the trip's departure time. The earliest possible walking start time of each request to the joint departure location of the shared trip is its requested departure time. For each request included in a shared trip, the arrival time at the desired arrival location must be within the maximum delay γ, taking into account a possible walking duration (see Constraint 7).

$$t_k^{dep} \leq \hat{t}_i^{dep} + \beta \qquad \forall\, i \in \mathcal{I}_k \;\; \forall\, k \in K \tag{4}$$

$$t_{k,v}^{taxi} + \theta_d(l_{k,v}^{taxi}, l_k^{dep}) \leq \hat{t}_i^{dep} + \beta \qquad \forall\, i \in \mathcal{I}_k \;\; \forall\, k \in K \;\; \forall\, v \in V \tag{5}$$

$$\hat{t}_i^{dep} + \theta_w(\hat{l}_i^{dep}, l_k^{dep}) \leq t_k^{dep} \qquad \forall\, i \in \mathcal{I}_k \;\; \forall\, k \in K \tag{6}$$

$$t_k^{arr} + \theta_w(l_k^{arr}, \hat{l}_i^{arr}) \leq \hat{t}_i^{arr} + \gamma \qquad \forall\, i \in \mathcal{I}_k \;\; \forall\, k \in K \tag{7}$$

Spatial constraints. The spatial constraints closely relate to the temporal constraints. Thus, Constraints 8 and 9 limit the walking distance of each request from the requested departure location and to the requested arrival location, respectively. The values for δ^{dep} and δ^{arr} are restricted by the maximum waiting time and the maximum delay, assuming a constant walking speed for the walking distance measured by the function Δ_w. Constraints 10 and 11 ensure that the driving distance, measured by the function Δ_d, of trip k is at least w times longer than the walking distance of each request and at least as long as a minimum distance g, respectively.

$$\Delta_w(\hat{l}_i^{dep}, l_k^{dep}) \leq \delta^{dep} \qquad \forall\, i \in \mathcal{I}_k;\ \ \forall\, k \in K \tag{8}$$

$$\Delta_w(l_k^{arr}, \hat{l}_i^{arr}) \leq \delta^{arr} \qquad \forall\, i \in \mathcal{I}_k;\ \ \forall\, k \in K \tag{9}$$

$$\frac{\Delta_w(\hat{l}_i^{dep}, l_k^{dep}) + \Delta_w(l_k^{arr}, \hat{l}_i^{arr})}{\Delta_d(l_k^{dep}, l_k^{arr})} \leq w \qquad \forall\, i \in \mathcal{I}_k;\ \ \forall\, k \in K \tag{10}$$

$$\Delta_d(\hat{l}_i^{dep}, \hat{l}_i^{arr}) \geq g \qquad \forall\, i \in \mathcal{I}_k;\ \ \forall\, k \in K \tag{11}$$

Cost constraints. Finally, we assume that customers are only willing to share a taxi with other customers if the shared trip leads to a cost saving of at least μ compared to a traditional taxi trip, as indicated by Constraint 12.

$$\hat{f}_i - f_k \geq \mu \qquad \forall\, i \in \mathcal{I}_k;\ \ \forall\, k \in K \tag{12}$$

4.3. Integer linear program

The mathematical model for optimizing one planning batch h is based on the idea of set partitioning. The input for the integer linear model is a set of due customer requests I_h^{due}, which would be too late to decide upon in the following planning batch $h+1$. The set of all feasible taxi trips is denoted by K_h. Typically, a single request i may be satisfied by more than one taxi trip k (e.g., as traditional trip or in a trip shared with other requests) leading to the number of feasible taxi trips exceeding the number of customer requests in a considered planning batch h. The binary decision variable y_i indicates whether a request i is rejected ($y_i = 1$) or not ($y_=0$). We use the binary decision variable $x_{v,k}$ (Eq. 16) to indicate whether taxi trip k executed by vehicle v is selected for service ($x_{v,k} = 1$) or not ($x_{v,k} = 0$). Since each taxi trip has a departure time and a departure and arrival location, selecting a trip implies deciding about the temporal and spatial conditions for answering a request.

$$\min \sum_{v\in V_h} \sum_{k\in \mathcal{K'}_{v,h}} cost(v,k) \cdot x_{v,k} + \sum_{i \in I_h^{due}} e \cdot y_i \tag{13}$$

$$s.t. \sum_{v\in V_h} \sum_{k\in \mathcal{K}_{i,v,h}} x_{v,k} + y_i = 1 \qquad \forall\, i \in I_h^{due} \tag{14}$$

$$\sum_{k\in \mathcal{K'}_{v,h}} x_{v,k} \leq 1 \qquad \forall\, v \in V_h \tag{15}$$

$$x_{v,k} \in \{0,1\} \qquad \forall\, k \in K_h \tag{16}$$

$$y_i \in \{0,1\} \qquad \forall\, i \in I_h^{due} \tag{17}$$

The objective is to minimize the sum of costs of the selected vehicle-trip-assignment plus a penalization e for each rejected request (Eq. 13). We consider objectives from both customers' and operators' perspectives to ensure the highest possible acceptance on both sides and achieve realistic outcomes. Additionally, this is achieved by considering the customer's perspective, e.g. in terms of walking-driving ratio, as hard constraints for feasible shared trips (compare Section 4.2) ensuring a specific service level for customers.

Constraint (14) assures that each due customer request $i \in I_h^{due}$ in planning batch h is either rejected or assigned to exactly one taxi trip $k \in K_h$, obtained through the row-wise sum of decision variables x_{vk}. The set $\mathcal{K}_{i,v,h}$ contains only the taxi trips k which include request i executed by vehicle v. Similarly, the set $\mathcal{K'}_{v,h}$ contains all taxi trips k which are executed by vehicle v. Constraint 15 guarantees that each available vehicle is assigned to at most one taxi trip. Finally, Constraints 16 and 17 define both decision variables to be binary.

4.4. Rolling horizon framework

As previously mentioned, we split an entire planning period (such as a day) into many deterministic planning batches h, i.e., we apply a rolling horizon approach. Each of these batches is characterized by a start time t_h^{start}. The optimization interval ω specifies the time until the start time of the subsequent batch t_{h+1}^{start} and thus, determines the available time for collecting newly available requests. Moreover, the time required to execute the optimization of one planning batch is considered by optimization duration ϵ. The temporal properties of the system, i.e., the sequence of several planning batches, are visualized in Figure 3.

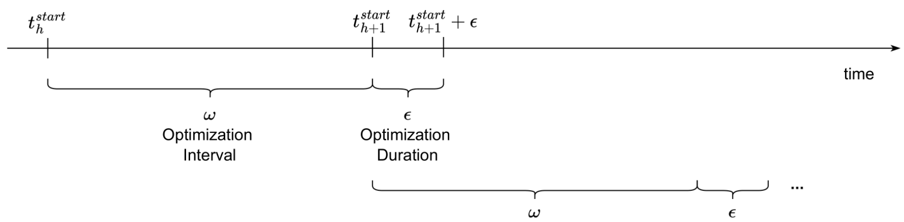

Figure 3: Temporal parameters of a planning batch h

In a planning batch, we only consider requests which would be too late to decide upon in the following planning batch. Thus, the response to a request is made as late as possible to increase the time for finding possible sharing partners.

An overview of the notation for all presented components can be found in Table 1.

Table 1: Notation overview

Notation	Definition
Indices	
$i \in I$	Request
$k \in K$	Taxi trip
$v \in V$	Taxi vehicle
$h \in H$	Planning batch
Decision Variables	
x_k	Taxi trip k is executed (=1) or not (=0)
y_i	Request i is rejected (=1) or not (=0)
Sets	
I_h^{due}	Due requests in planning batch h
K_h	Taxi trips in planning batch h
V_h	Set of taxi vehicles available in planning batch h
$\mathcal{I}_k$	Set of requests served by taxi trip k
$\mathcal{K}_{i,v,h}$	Taxi trips in planning batch h that include request i and can be executed by vehicle v
$\mathcal{K'}_{v,h}$	Taxi trips in planning batch h that can be executed by vehicle v
Parameters	
$\hat{t}_i^{req}$	Request time of request i
$\hat{t}_i^{not}$	Latest accepted notification time of request i
$\hat{t}_i^{dep}$	Requested departure time of request i
$\hat{t}_i^{arr}$	Expected arrival time of request i
$\hat{l}_i^{dep}$	Requested departure location of request i
$\hat{l}_i^{arr}$	Requested arrival location of request i

$\hat{p}_i$	Requested number of passengers of request i
$\hat{f}_i$	Fee for customer request i for a traditional taxi trip
λ_i	Lead time of request i
α	Minimum notification time of a request
β	Maximum waiting time of a request i
γ	Maximum delay of a request
δ_i^{dep}	Maximum walking distance of request i before the departure
δ_i^{arr}	Maximum walking distance of request i after the arrival
t_k^{dep}	Departure time of taxi trip k
t_k^{arr}	Arrival time of taxi trip k
l_k^{dep}	Departure location of taxi trip k
l_k^{arr}	Arrival location of taxi trip k
p_k	Number of passengers of taxi trip k
f_k	Fee per request for taxi trip k
c_k	Costs of taxi trip k executed by vehicle v
$t_{v,k}^{taxi}$	Time of earliest availability of taxi v to drive to the departure of trip k
$l_{v,k}^{taxi}$	Location of taxi v before driving to the departure of trip k
t_h^{start}	Start time of planning batch h
ω	Optimization interval of a planning batch
ϵ	Optimization duration of a planning batch
r^{max}	Maximum number of requests per shared taxi trip
q	Seating capacity of each taxi in the vehicle fleet
w	Walking-Driving ratio
μ	Minimum cost savings per request as prerequisite for sharing
g	Minimum distance of a request to be considered
u_b	Unitary value of waiting
u_w	Unitary value of walking
u_d	Unitary value of driving
u_o	Unitary cost of operating a vehicle

e	Penalty expense for rejecting a request

5. Computational study

In this section, we describe our computational study by first presenting the experimental design and then reporting the computational results obtained.

5.1. Experimental design

In our experiments, we utilize taxi data collected in New York City (NYC City Taxi & Limousine Commission: TLC, 2017) and Porto (Moreira-Matias et al., 2013). Both data sets have been used in other studies on TRS systems (see e.g., Barann et al., 2017; d'Orey et al., 2012) and represent traditional taxi systems without any sharing. We draw on these two data sets, as they have different spatio-temporal distributions, i.e., the number of daily requests in New York City is around 100 times higher than in Porto, while the urban area of New York City is only 20 times larger than the area of Porto. In absolute numbers, there are on average around 400,000 requests in New York City and 4,000 in Porto. This allows us to evaluate the potential of the system more broadly. The data includes the origin/destination and the number of passengers of taxi trips, as well as the time the trips started/ended and the costs of the trips. However, the data does not contain lead times but solely the time the trips physically started and therefore, need to be imputed. Previous studies often used a constant lead time (Agatz et al., 2011; Stiglic et al., 2018). However, we assume that a Poisson distribution gives more realistic estimations of lead times due to its stochastic nature. Furthermore, it has a lower bound at a value of 0 and might therefore result in more realistic estimations compared to other distributions. Therefore, we draw from a Poisson distribution with rate λ and subtract the drawn values from the start time of the trip. This is done two times with different random seed values.

We consider a 24-hour planning horizon starting at 0:00 (midnight) on Tuesday, April 8th, 2014. For New York City, our analysis is limited to taxi requests with both pickup and drop-off locations confined within Manhattan; consequently, throughout this paper, the terms 'NYC' and 'Manhattan' are used interchangeably. To approximate the distance between two points, we apply the Haversine distance and multiply it by the correction factor of 1.5, which is a rough approximation of a typical shortest path distance on a road network (Ehmke and Campbell, 2014). We note that this factor is likely different in Manhattan but alternatives such as the Manhattan distance would give overly optimistic approximations. Therefore, we assume a constant factor over all regions. Further, we assume a car capacity q of 3 passenger seats, due to reasons of passenger comfort. We've set the maximum number of requests per shared taxi trip to the same value ($r^{max} = 3$), ensuring manageable computational complexity. Moreover, we assume a constant walking speed of 5.1 km/h (which is roughly the average normal walking speed of people (Bohannon and Andrews, 2011)) that is independent of time and location. The minimum distance of a request to be considered (g) is set to 425 meters. This is the distance that a person would walk in 5 minutes (assuming the previously mentioned constant walking speed). The walking-driving ratio (w) is defined at 0.2. This ensures that with a constant driving speed of 30 km/h for all taxis, the maximum possible walking duration remains within the range of the driving duration of the shared trip. As sharing surcharge we assume \$2.50 (Barann et al., 2017). Inspired by the approach of Fielbaum et al. (2021), we consider $u_b = u_w = 2u_d$ and $u_o = 1.5u_d$ with $u_d = 1$ for calculating the costs associated with a vehicle-trip-assignment as described in formula (1). The penalty costs e for rejecting a request are set to one million. This high value is chosen such that requests are only rejected when a customer is not reachable or no taxi is available.

For New York City, a fleet size of 5000 vehicles is derived from one-third of the total taxi licenses in the city (NYC City Taxi & Limousine Commission: TLC, 2014), representing the available taxis when assuming three shifts per day. For Porto, guided by Moreira-Matias et al. (2012), we've selected a fleet size of

400 vehicles. Inspired by Bertsimas et al. (2019), we have set the initial vehicle distribution at random locations, which follow the distribution of customer drop-off locations in NYC and Porto right before the start of the planning horizon.

To investigate the system's sensitivity, we vary parameters of two types, which we refer to as exogenous and system design parameters. *Exogenous* parameters are set outside the TRS system, i.e., they specify the system's input and are not within the decision power of the system provider, such as the volume of customer requests. In contrast, *system design* parameters determine the system design and must be fixed by the system provider, and include, for example, the optimization frequency. For the sensitivity study, we focus on varying three central exogenous parameters, which influence the system's input, i.e., the customer requests, significantly: City, density of requests, and announcement lead time (with parameter λ representing the respective distribution). The varied system design parameters consist of the three most important parameters for ensuring a promised level of service to customers, i.e., minimum notification time α, maximum waiting time β and maximum walking distance ψ, and one parameter concerning the rolling horizon approach, i.e., optimization interval ω. The values of these system design and exogenous parameters are presented in Table 2. We create a scenario for each possible combination of these parameter values.

To explore diverse demand scenarios, we manipulate request density ϕ by randomly filtering out customer requests, examining three cases: 50%, 75%, and 100% of the requests from the respective data sets. For the value of λ, we orientate ourselves on the constant value chosen by Agatz et al. (2011) and Stiglic et al. (2018), and choose settings where customer requests arrive spontaneously (900 seconds) and with higher lead times (1,800 seconds and 3,600 seconds). For the minimum notification time α, we consider three settings: In the first setting (0), customers only need to be notified at their requested departure time. In the second and third setting, customers need to be notified at least 3 or 5 minutes before they start their travel. We perceive these values as reasonable because previous works did not consider a notification time,

Table 2: Parameter settings for scenarios

	Parameter	**Values**
exogenous	City η	New York City; Porto
	Density ϕ (in %)	50; 75; 100
	Lead time parameter λ (in s)	900; 1,800; 3,600
system design parameters	Minimum notification time α (in s)	0; 180; 300
	Maximum waiting time β (in s)	300; 600; 900
	Maximum walking distance δ (in m)	425; 850; 1275
	Optimization interval ω (in s)	60; 180; 300

i.e., implicitly assumed a value of 0 (see e.g., Wang et al., 2018). Concerning the maximum waiting time β, similar to Barann et al. (2017) and Qian et al. (2017), we consider a scenario where customers can only wait shortly (300 seconds), moderately (600 seconds) and one in which customers are highly flexible (900 seconds). Furthermore, the parameter that determines the maximum delay (γ) is set to two times the maximum waiting time. The maximum walking distance δ is varied independently of the maximum waiting time, although it is additionally constrained by it. Accordingly, we select the values of 425, 850, and 1275 meters. Finally, the optimization interval ω is set to 60, 180, and 300 seconds. These values are based on the studies of Kucharski and Cats (2020) and Fielbaum et al. (2021). We perform an exhaustive search over the space of possible parameter combinations, resulting in a total of 2,916 instances.

The experiments run on compute nodes with 3.5 GHz and 256 GB RAM within Paderborn University's high-performance computer cluster, allowing parallel execution of different scenarios. Furthermore, we use the Gurobi 10 solver with standard settings and a time limit equal to a constant optimization duration ϵ of 60 seconds. Pretests have shown the solver consistently converges fast towards the optimal solution, i.e., optimality gaps are consistently near 0%. The max gap in New York is 1.1% and all instances are optimally solved in Porto. Further summary statistics are to be found in the Appendix A (Table A.8).

5.2. Results

In this subsection, we present the results of the conducted experiments. The results are then discussed and interpreted in Section 6. Table 3 displays descriptive statistics for the rejection rate and sharing rate across various scenarios in New York City and Porto. Additionally, the table includes the absolute count of responded requests, consisting of both accepted and rejected requests about which customers have been informed within the considered planning horizon. The rejection and sharing rate represent the proportions of rejected and shared requests among all responded requests, enabling the derivation of the absolute count of shared and rejected requests from the data in Table 3. The number of responded requests for the considered planning horizon varies due to different values for density, the optimization interval, and the notification time.

Table 3: Summary statistics for rejection rate and sharing rate

	Rejection rate (%)		Sharing rate (%)		Responded requests	
	NYC	Porto	NYC	Porto	NYC	Porto
Mean	3.24	1.01	6.05	0.28	239421	2775
Median	1.40	0.00	5.38	0.27	240343	2780
Min	0.00	0.00	1.32	0.11	158788	1829
Max	22.20	16.45	16.87	0.73	322347	3736
Stdv	4.21	2.30	3.44	0.13	65222	755

Table 4 provides descriptive statistics for metrics describing the service level of customers, specifically, the average driving time, average walking time, and average waiting time of resulting trips.

To provide more detailed insights into the sharing potential and the rejection rate, Figure 4 and 5 exhibit boxplots of the distribution of the sharing rate and

Table 4: Summary statistics for average driving, walking and waiting times

	Average driving time (sec)		Average walking time* (sec)		Average waiting time (sec)	
	NYC	Porto	NYC	Porto	NYC	Porto
Mean	387	599	74	10	256	50
Median	388	600	65	8	228	24
Min	367	588	57	4	126	11
Max	397	607	118	65	575	256
Stdv	6	4	14	6	97	50

*only the walking time to the departure location is considered (as in the objective function)

the rejection rate in New York City, respectively, split by density ϕ. Given the limited vehicle fleet, there exists a reciprocal impact between the rejection rate and the sharing rate. Increased ride sharing enables the fulfillment of more requests and subsequently reduces the number of rejections due to vehicle unavailability. For NYC, this relationship is represented in Figure 6, depicting the rejection rate on the x-axis and the sharing rate on the y-axis. Respective Figures for Porto can be found in Appendix A (Figure A.7 to A.9).

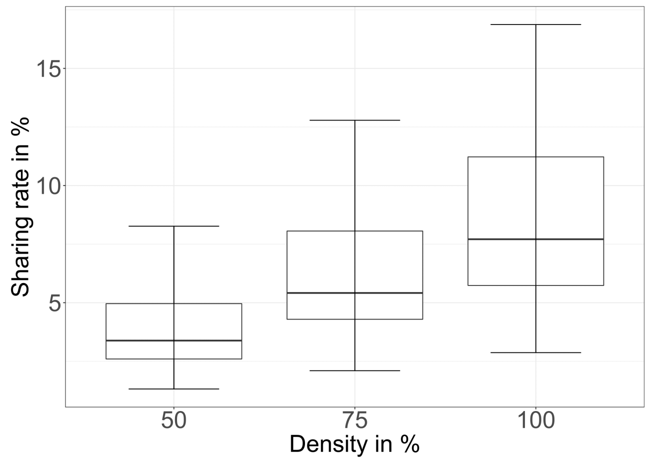

Figure 4: Boxplot showing sharing rate (in %) in Manhattan

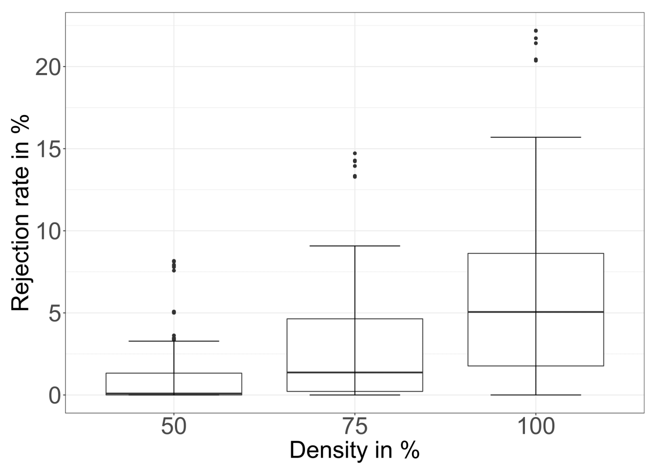

Figure 5: Boxplot showing rejection rate (in %) in Manhattan

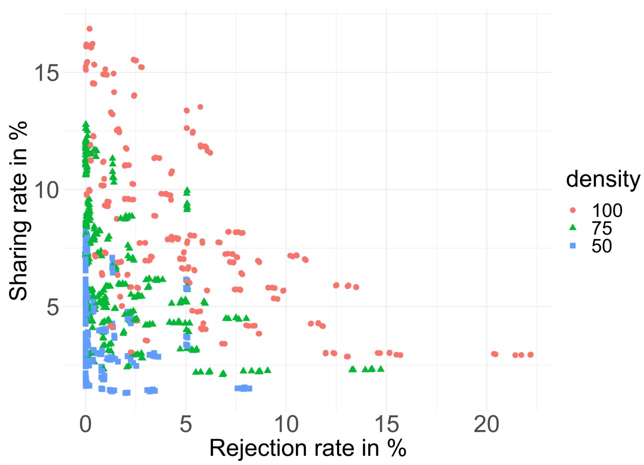

Figure 6: Relation between sharing and rejection rate in Manhattan

Finally, we conduct ordinary least square regressions using the aforementioned metrics (rejection rate, sharing rate, average driving time, average walking time, and average waiting time) as dependent variables. The aim is to quantify the impact of system design and exogenous parameters, as outlined in Section 5.1, on the system's efficiency. Within this model, we introduce interaction terms by associating the city parameter with all parameters specified in Section 5.1. This approach allows us to analyze how the coefficients vary across different cities. Given that the explanatory variables represent only 'low', 'medium', and 'high' scenarios, we treat them as categorical variables in the model, where the medium-value is the base value for the regression. This approach results in a total of five regressions in the form of Equation 18.

$$\begin{aligned} y = b_0 + \eta \cdot b_1 + \phi \cdot b_2 + \lambda \cdot b_3 + \alpha \cdot b_4 + \beta \cdot b_5 + \delta \cdot b_6 + \omega \cdot b_7 + \\ \eta\phi \cdot b_8 + \eta\lambda \cdot b_9 + \eta\alpha \cdot b_{10} + \eta\beta \cdot b_{11} + \eta\delta \cdot b_{12} + \eta\omega \cdot b_{13} + error \end{aligned} \tag{18}$$

The results of these regressions are presented in Table 5 and 6. Since NYC is the base value for the 'city'-variable, all coefficients above the dashed line describe the scenarios in NYC. To interpret scenarios in Porto, the values below the dashed line need to be additionally considered.

Table 5: Regression models for rejection rate and sharing rate

	Rejection rate (in %)	Sharing rate (in %)
(Intercept)	0.61^{**}	7.87^{***}
density50	-1.64^{***}	-2.28^{***}
density100	3.23^{***}	2.20^{***}
leadTimeLambda15	1.09^{***}	-0.46^{***}
leadTimeLambda60	-0.18	-0.05
minNotTime0	1.56^{***}	0.43^{***}
minNotTime300	-0.12	-0.33^{***}
maxWaitingTime300	4.01^{***}	-2.77^{***}
maxWaitingTime900	-1.47^{***}	0.31^{***}
maxWalkingDist425	0.21	-1.96^{***}
maxWalkingDist1275	0.06	0.14^{*}
optInterval60	2.20^{***}	-2.76^{***}
optInterval300	-1.01^{***}	2.12^{***}
cityPorto	-1.47^{***}	-7.65^{***}
cityPorto:density50	1.26^{***}	2.20^{***}
cityPorto:density100	-2.34^{***}	-2.08^{***}
cityPorto:leadTimeLambda15	0.13	0.47^{***}
cityPorto:leadTimeLambda60	0.23	0.12
cityPorto:minNotTime0	-0.56^{**}	-0.38^{***}
cityPorto:minNotTime300	0.68^{***}	0.35^{***}
cityPorto:maxWaitingTime300	-2.25^{***}	2.74^{***}
cityPorto:maxWaitingTime900	1.42^{***}	-0.30^{***}
cityPorto:maxWalkingDist425	-0.21	1.95^{***}
cityPorto:maxWalkingDist1275	-0.06	-0.13
cityPorto:optInterval60	-1.83^{***}	2.70^{***}
cityPorto:optInterval300	1.25^{***}	-2.05^{***}
R^2	0.64	0.95
Adj. R^2	0.63	0.95
Num. obs.	2916	2916

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 6: Regression models for avg. driving-, walking- and waiting time

	Avg. driving time (in sec)	Avg. walking time (sec)	Avg. waiting time (sec)
(Intercept)	387.06***	84.83***	289.82***
density50	4.65***	−3.84***	−65.26***
density100	−6.33***	2.99***	66.76***
leadTimeLambda15	0.52*	0.27	0.70
leadTimeLambda60	−0.99***	−0.35	−5.40*
minNotTime0	−2.86***	2.21***	29.80***
minNotTime300	1.37***	−1.23**	−18.17***
maxWaitingTime300	−2.78***	−18.44***	−113.42***
maxWaitingTime900	0.53*	0.10	49.97***
maxWalkingDist425	3.47***	−16.03***	−20.47***
maxWalkingDist1275	−0.68**	0.93*	−1.20
optInterval60	1.36***	1.97***	10.72***
optInterval300	0.65**	−1.35***	−32.26***
cityPorto	209.94***	−76.84***	−255.66***
cityPorto:density50	−2.98***	4.08***	43.20***
cityPorto:density100	7.19***	−2.97***	−29.82***
cityPorto:leadTimeLambda15	2.06***	−0.63	−16.65***
cityPorto:leadTimeLambda60	4.03***	0.29	−4.73
cityPorto:minNotTime0	1.21***	1.90***	18.97***
cityPorto:minNotTime300	−1.32***	3.15***	13.43***
cityPorto:maxWaitingTime300	1.32***	18.07***	98.00***
cityPorto:maxWaitingTime900	−0.50	0.01	−47.90***
cityPorto:maxWalkingDist425	−3.45***	15.22***	19.03***
cityPorto:maxWalkingDist1275	0.64*	−0.80	1.87
cityPorto:optInterval60	−1.03***	−3.65***	15.71***
cityPorto:optInterval300	−0.93**	5.33***	35.90***
R^2	1.00	0.97	0.93
Adj. R^2	1.00	0.97	0.93
Num. obs.	2916	2916	2916

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

6. Discussion

In this section, we discuss the results of our computational study by first interpreting the variations of rejection rate and sharing rate and then deriving design implications that need to be considered when designing a one-to-one TRS system.

6.1. Variations of rejection rate, sharing rate and service quality

There are large differences in the rejection rate and the sharing rate resulting from the variation of parameters, which determine the mobility environment and the design of the TRS system. Please notice that, as we regress rejection and sharing rates which are measured in percentage, we talk about percentage points when we say that sharing/rejection rates increase/decrease. For example, if the intercept is 5%, an increase of 10%-points indicates 15% and not 5.5%. Below, we will individually analyze the influence of varying each parameter listed in Table 2 starting with the exogenous parameters followed by the system design parameters.

City. From Table 3 we see that the mean rejection rate is 3.24% in New York and 1.01% in Porto. This difference is likely caused by the fact that the requests-per-taxi-ratio is higher in New York than in Porto (see Subsection 5.1 for a detailed explanation of our chosen taxi fleet size), i.e., a single taxi needs to serve fewer requests in Porto than in New York. The median values in both cities are lower than the mean, which indicates that in an average scenario, very few rejections occur. The mean sharing rate is 6.05% in New York and 0.28% in Porto. These results are also supported by the regression results (Table 5). The city of Porto variable has a significant negative coefficient for both rejection (-1.47) and sharing rates (-7.65). The relatively low sharing rates can be attributed to a few factors. First, the walking-driving ratio of 0.2 restricts a lot of potential sharing, as rides where customers need to walk considerably more compared to the taxi ride duration are deemed unfeasible. We are among the few to constrain this ratio, a factor also acknowledged in the approaches of Barann et al. (2017) and Dieter et al. (2023), and likely considered relevant by many customers. This constraint notably impacts Manhattan, where trip durations are relatively short, amplifying its effect (compare Table 4). Second, in the objective function, driving is less penalized than walking. This results in cases where requests are not shared even if no hard constraints are violated, provided enough taxis are available. These findings correspond to the service quality metrics presented in Table 4 and 6, displaying low values, particularly for the average walking time (on average 74 seconds in NYC), as well as for the waiting time (on average 256 seconds in NYC).

Density. The density parameter simulates different demand scenarios in the cities. It directly affects the requests-per-taxi-ratio. Thus, a higher density leads to increased rejection rates and sharing rates. In NYC, an increase in density from 75% to 100% corresponds to an increase in the average sharing rate by more than 2%-points and in the average rejection rate by more than 3%-points (see boxplots in Figure 4 and 5). The regression coefficients in Table 5 confirm these observations and indicate a contrasting effect when the density decreases to 50%, leading to lower rejection and sharing rates than those observed at a density of 75%. One reason for these comparatively high effects could be the absence of a rebalancing algorithm, which would redistribute empty taxis to regions with high demand. The direction of the effects is similar for Porto, but their effect size there is noticeably weaker, possibly due to the generally lower requests-per-taxi-ratio and the lower sharing potential. This is supported by the regression coefficients, which, while significant, exhibit lower values for Porto (e.g. $2.20 - 2.08 = 0.12$ percentage points for a density of 100%). Regarding service quality, the density of requests primarily influences the waiting time. According to the regression results, a 25% increase in request density results in a one-minute increase of waiting time.

Lead time parameter. The base value in our regression is set to the mid-value, i.e., customers request their taxi on average 30 minutes before their desired start time. From Table 5, we can see that a value of 15 minutes increases the rejection rate by 1.09% compared to the base case of 30 minutes. Further, this relation is significant with a p-value of less than 0.1%. We deduct that more requests on short notice lead to a larger amount of rejections. This is because in case the lead time of a customer is lower than the sum of optimization interval and optimization duration, the customer is directly rejected without being considered for optimization. When looking at the sharing rate, a parameter of 15 leads to 0.46%-points fewer sharing and the coefficient is also significant. An explanation for this observation is that longer lead times provide information earlier compared to shorter lead times. Consequently, with longer lead times, the system can make better-informed decisions, facilitating the identification of more suitable sharing partners. Conversely, shorter lead times may delay crucial information, potentially leading to less informed decisions in finding suitable sharing partners. When looking at the parameter of 60 minutes, the effects on rejection and sharing rate are both slightly negative and not significant. This implies that in contrast to 15 minutes, it does not make a big difference if the lead time is 30 or 60 minutes. This is likely because both these durations provide ample time for the system to make well-informed decisions. Another interesting observation is that the observed differences in the sharing rate between lead time values of 15 and 30 are less present in Porto, as the interaction effect between Porto and a lead time setting of 15 minutes has a positive and significant coefficient of 0.47, which equals out the coefficient for both Porto and New York (-0.46). A possible explanation for this observation is that almost no sharing occurs in Porto in any of the tested scenarios, disregarding the lead time parameter. Moreover, the lead time parameter demonstrates no substantial influence on service quality, as evidenced by the respective regression coefficients (Table 6).

Minimum notification time. For the minimum notification time, the baseline in our regression is set at the mid-value, where customers receive their notification at least 3 minutes before their desired start time. A change in the notification time, as indicated by the regression, only exhibits a significant influence on the rejection rate when the notification time is reduced (in this case, to 0 minutes). Unexpectedly, despite providing the system additional 3 minutes to decide on feasible vehicle-trip-assignments, reducing the notification time leads to an increase in the rejection rate (by 1.56%-points in NYC). However, with the elimination of the notification time, taxis have less time available to reach the pick-up location promptly, leading to higher rejection rates and higher average waiting times (e.g. 29.8 seconds in NYC in Table 4). At the same time, the impact of the notification time on the sharing rate follows the expected pattern: reducing the notification time increases the sharing rate (by 0.43%-points in NYC), whereas increasing the notification time decreases it (by -0.33%-points in NYC). This relationship is attributed to the system having more time to find sharing partners, leading to more shared trips. Again, similar relationships can be observed for Porto, albeit to a lesser extent.

Maximum waiting time. The system design parameter with the highest impact on the rejection rate is the maximum waiting time β. While the regression baseline assumes a maximum waiting time of 10 minutes, it suggests that reducing the parameter by 5 minutes leads to an increase in the rejection rate of 4.01%-points, while increasing the parameter by 5 minutes results in a decrease in the rejection rate of $-1,47$%-points in NYC. The increasing rejections with shorter waiting times stem from two main reasons: Firstly, the limited availability of taxis unable to reach the pick-up location within the abbreviated time frame. Secondly, and more crucially, the reduced potential for shared trips, which leads to fewer available vehicles even with relatively low demand. This is supported by the regression coefficients for the model with the sharing rate as dependent variable, which show a decrease of -2.77%-points for $\beta = 300$ and an increase of 0.31%-points for $\beta = 900$ in NYC. This relationship is depicted in Figure 6. The results positioned towards the bottom-right corner correspond to lower values of the maximum waiting time, whereas the results associated with a high maximum waiting time are represented by the points situated towards the top-left. In Porto, this correlation is somewhat less prominent, as evidenced by the lower regression coefficients (see also Figure A.9). Similar to other parameters, service quality is primarily influenced by changes in waiting time. Although it increases or decreases with enlargement or reduction of the maximum waiting time parameter, it does so without fully exploiting the boundaries. According to the regression, in NYC, reducing the maximum waiting time to 5 minutes results in a decrease in the average waiting time of about 2 minutes, whereas increasing the maximum waiting time to 15 minutes leads to an increase in the average waiting time of less than 1 minute.

Maximum walking distance. The regression indicates that a change in the maximum walking distance has no significant impact on the rejection rate. In terms of sharing rate, a noteworthy impact is observed solely when the maximum walking distance is decreased to 425 meters, leading to a decrease of 1.96%-points in the sharing rate for NYC. The explanation for the limited influence of the walking distance is its restriction by the waiting distance. Given that the regression's baseline for the maximum waiting time is set at 10 minutes, only a maximum walking distance of 425 meters imposes a stronger walking constraint, thus resulting in the observation of a significant coefficient in this instance. These effects diminish in the case of Porto because of the few total shared trips. The average walking time remains almost unaffected by changes in the maximum waiting time (see e.g. the coefficient of -16.03 for maxWalkingDist425).

Optimization interval. The optimization interval is the system design parameter with the highest impact on the sharing rate. According to Table 5, for NYC, it's evident that a value of 60 seconds decreases the sharing rate by -2.76%-points compared to the base case of 180 seconds. Conversely, a value of 300 seconds increases the sharing rate by 2.12%-points compared to the base case of 180 seconds. Additionally, this relationship is statistically significant with a p-value of less than 0.1%. Our results are consistent with findings from the literature. For example, Kucharski and Cats (2020) noted that longer scheduling horizons foster optimal trip-sharing, while shorter horizons hinder this potential. The authors observe that a 5-minute interval is enough for customers to optimally share their trips. For a short optimization interval, the declining sharing potential accounts for the rising rejection rate (2.20%-points for NYC), whereas for a larger optimization interval, the increasing sharing potential explains the decreasing rejection rate (-1.01%-points for NYC). Increasing the optimization interval to 5 minutes even has a positive effect on the average waiting time, reducing it by 30 seconds for NYC (Table 6).

6.2. Design implications

With our experiments, we not only demonstrate the instantiation and application of the suggested TRS system but also point to general issues that need to be considered by organizational decision-makers when implementing the proposed system in practice. We do this by drawing on the regression results with the parameters in Table 2 as independent variables. While the exogenous parameters serve to assess whether implementing a TRS system is viable, the system design parameters reflect the design choices available for the TRS system. The directions of effects of these parameters on the system's performance, in terms of rejection rate, sharing rate and service quality, are similar for both cities, New York City and Porto.

It's essential, as a primary consideration before implementing a one-to-one taxi ridesharing system, to ascertain whether the city, where the system is intended to be set up, has sufficient ride requests with sharing potential. Such a system may not be justified in a city like Porto with a relatively low volume of daily requests that are geographically dispersed. As our experiments have demonstrated, a high request density significantly affects rejection and sharing rates of one-to-one TRS, especially evident in NYC. This emphasizes the necessity of tailored strategies to manage demand fluctuations in densely populated areas, requiring the implementation of rebalancing algorithms and dynamic fleet size adjustments. Moreover, the system operator should encourage users to make early enough requests, e.g., by incentivizing early requests. However, lead times don't necessarily need to be excessively long. Beyond a certain threshold, extended lead times may not necessarily result in increased sharing rates and decreased rejection rates.

If the exogenous parameters create favorable conditions for a one-to-one TRS system, the system's design becomes crucial in ensuring good performance. Table 7 summarizes the essential design implications crucial for optimizing one-to-one TRS systems.

7. Conclusions

In this work, we address the design of TRS with shared pick-up and drop-off locations for sustainable shared mobility as an important research area in modern transportation. While the literature on shared mobility shows consensus on the large potential of TRS, it has been remarkably silent on how to design such systems. However, getting insights into such design issues is crucial for developing sustainable TRS systems.

We approach the above-mentioned research gap with a computational study and provide several contributions: (i) We propose a mathematical formulation of the one-to-one TRS problem; (ii) we suggest the application and describe the algorithmic implementation of a rolling horizon approach to account for the dynamic problem nature; (iii), we conduct extensive computational experiments based upon real-world data of the cities of New York and Porto, including a sensitivity analysis of various system design and exogenous parameters; and (iv) we derive prescriptive managerial implications for practitioners who plan to design and implement one-to-one TRS systems. These implications are based upon our findings that the maximum waiting time of customers has the strongest leverage effect on rejection and sharing rates. Managerial implications are also derived from our findings that some other system design parameters have a significant impact, such as the optimization interval and the minimum time customers need to be notified before their requested departure time (minimum notification time).

Our study opens avenues for a variety of future research directions: While we assume customer behavior to be homogeneous, future research may distin-

Table 7: Design implications

No.	Implication
1	The *maximum waiting time* is the primary factor influencing rejection rates and ride-sharing potential in one-to-one TRS systems. Operators are strongly advised to implement market mechanisms that increase customer flexibility regarding waiting times. Thoughtful management of this parameter is vital: reducing it below 10 minutes may significantly increase rejections, while increasing it could enhance shared trips but necessitates careful balance for optimal service quality.
2	The size of the *optimization interval* significantly influences the sharing rate in TRS systems during a rolling horizon approach. Operators are advised to avoid excessively short intervals and aim for approximately 5-minute durations, as this can enhance sharing rates, thereby potentially reducing rejection rates.
3	The *minimum notification time* influences rejection rates by affecting prompt taxi arrivals, while extending it reduces the sharing rate. We recommend operators to consider informing customers shortly before the requested departure, such as 3 minutes, ensuring customers' planning security and allowing taxis ample time to reach the pick-up location.
4	The *maximum walking distance* shows no substantial impact on rejection rates but affects sharing rates, notably observed when reduced to distances below what is reachable within the maximum waiting time. Therefore, operators of TRS systems are recommended to consider the maximum walking distance carefully, ensuring it aligns with distances reachable within the maximum waiting time.

guish the willingness to share taxis at an individual level. This extension might include incentivizing customers when they, for example, accept higher walking distances or longer lead times for their requests. One option to follow this research direction involves dynamic pricing or other incentive schemes, such as future discounts. Second, repositioning of idle taxis can further increase the efficiency of the system and is therefore worthwhile studying. Third, in order to increase the positive environmental impact of the system, the operator could prioritize long taxi trips when dispatching taxis as short trips are more likely to be substituted by more sustainable means of transportation, such as walking or bike sharing. Fourth, and related to the previous research direction, other transportation options, such as scooter or bike-sharing systems, may be included in the TRS system. For example, in case a customer i needs to walk a long distance, a scooter may be reserved for customer i, which allows the operator of the TRS system to assign customer i to a group of customers joining each other at a location that would not have been reachable by customer i only by walking.

Acknowledgements

The authors gratefully acknowledge the computing time provided to them on the high-performance computers Noctua 2 at the NHR Center PC2. These are funded by the Federal Ministry of Education and Research and the state governments participating on the basis of the resolutions of the GWK for the national highperformance computing at universities (www.nhr-verein.de/unsere-partner).

Declarations of interest

None.

References

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

Agatz, N., Erera, A.L., Savelsbergh, M.W., Wang, X., 2011. Dynamic ride-sharing: A simulation study in metro Atlanta. Procedia-Social and Behavioral Sciences 17, 532–550. doi:10.1016/j.trb.2011.05.017.

Aissat, K., Oulamara, A., 2014. A priori approach of real-time ridesharing problem with intermediate meeting locations. Journal of Artificial Intelligence and Soft Computing Research 4, 287–299.

Aliari, S., Haghani, A., 2022. Alternative pickup locations in taxi-sharing: A feasibility study. Transportation Research Record , 03611981221104690doi:10.1177/03611981221104690.

Alonso-Mora, J., Samaranayake, S., Wallar, A., Frazzoli, E., Rus, D., 2017. On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. Proceedings of the National Academy of Sciences 114, 462–467. doi:10.1073/pnas.1611675114.

Aïvodji, U.M., Gambs, S., Huguet, M.J., Killijian, M.O., 2016. Meeting points in ridesharing: A privacy-preserving approach. Transportation Research Part C: Emerging Technologies 72, 239–253. doi:10.1016/j.trc.2016.09.017.

Balardino, A.F., Santos, A.G., 2016. Heuristic and exact approach for the close enough ridematching problem, in: Hybrid Intelligent Systems: 15th International Conference HIS 2015 on Hybrid Intelligent Systems, Seoul, South Korea, November 16-18, 2015 15, Springer. pp. 281–293.

Barann, B., Beverungen, D., Müller, O., 2017. An open-data approach for quantifying the potential of taxi ridesharing. Decision Support Systems 99, 86–95. doi:10.1016/j.dss.2017.05.008.

Bathla, K., Raychoudhury, V., Saxena, D., Kshemkalyani, A.D., 2018. Real-time distributed taxi ride sharing, in: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), IEEE. pp. 2044–2051. doi:10.1109/ITSC.2018.8569315.

Bertsimas, D., Jaillet, P., Martin, S., 2019. Online vehicle routing: The edge of optimization in large-scale applications. Operations Research 67, 143–162.

Bohannon, R., Andrews, A., 2011. Normal walking speed: A descriptive meta-analysis. Physiotherapy 97, 182–9. doi:10.1016/j.physio.2010.12.004.

Bruck, B.P., Incerti, V., Iori, M., Vignoli, M., 2017. Minimizing CO2 emissions in a practical daily carpooling problem. Computers & Operations Research 81, 40–50. doi:10.1016/j.cor.2016.12.003.

Chen, W., Mes, M., Schutten, M., Quint, J., 2019. A ride-sharing problem with meeting points and return restrictions. Transportation science 53, 401–426.

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

Correia, G., Viegas, J.M., 2009. A conceptual model for carpooling systems simulation. Journal of Simulation 3, 61–68. doi:10.1057/jos.2008.4.

Czioska, P., Kutadinata, R., Trifunović, A., Winter, S., Sester, M., Friedrich, B., 2019. Real-world meeting points for shared demand-responsive transportation systems. Public Transport 11, 341–377.

Czioska, P., Trifunović, A., Dennisen, S., Sester, M., 2017. Location-and time-dependent meeting point recommendations for shared interurban rides. Journal of location based services 11, 181–203.

Demissie, M.G., Kattan, L., Phithakkitnukoon, S., de Almeida Correia, G.H., Veloso, M., Bento, C., 2020. Modeling location choice of taxi drivers for passenger pickup using gps data. IEEE Intelligent Transportation Systems Magazine 13, 70–90. doi:10.1109/MITS.2020.3014099.

Dieter, P., Stumpe, M., Ulmer, M.W., Schryen, G., 2023. Anticipatory assignment of passengers to meeting points for taxi-ridesharing. Transportation Research Part D: Transport and Environment 121, 103832.

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

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.

Engelhardt, R., Bogenberger, K., 2021. Benefits of flexible boarding locations in on-demand ride-pooling systems, in: 2021 7th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), IEEE. pp. 1–6. doi:10.1109/MT-ITS49943.2021.9529284.

Fielbaum, A., Bai, X., Alonso-Mora, J., 2021. On-demand ridesharing with optimized pick-up and drop-off walking locations. Transportation Research Part C: emerging technologies 126, 103061. doi:10.1016/j.trc.2021.103061.

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

Gökay, S., Heuvels, A., Krempels, K.H., 2019. On-demand ride-sharing services with meeting points., in: VEHITS, pp. 117–125.

Ham, A., 2021. Dial-a-ride problem with meeting point feature known-as express-pool. IEEE Access 9, 86404–86411. doi:10.1109/ACCESS.2021.3089275.

Hosni, H., Naoum-Sawaya, J., Artail, H., 2014. The shared-taxi problem: Formulation and solution methods. Transportation Research Part B: Methodological 70, 303–318. doi:10.1016/j.trb.2014.09.011.

Jung, J., Jayakrishnan, R., Park, J.Y., 2016. Dynamic shared-taxi dispatch algorithm with hybrid-simulated annealing. Computer-Aided Civil and Infrastructure Engineering 31, 275–291. doi:10.1111/mice.12157.

Khazbak, Y., Fan, J., Zhu, S., Cao, G., 2020. Preserving personalized location privacy in ride-hailing service. Tsinghua Science and Technology 25, 743–757. doi:10.26599/TST.2020.9010010.

Kucharski, R., Cats, O., 2020. Exact matching of attractive shared rides (exmas) for system-wide strategic evaluations. Transportation Research Part B: Methodological 139, 285–310. doi:10.1016/j.trb.2020.06.006.

Li, N., Kong, L., Shu, W., Wu, M.Y., 2019. Benefits of short-distance walking and fast-route scheduling in public vehicle service. IEEE Transactions on Intelligent Transportation Systems 21, 3706–3717.

Li, W., Cao, J., Guan, J., Zhou, S., Liang, G., So, W.K., Szczecinski, M., 2018a. A general framework for unmet demand prediction in on-demand transport services. IEEE Transactions on Intelligent Transportation Systems 20, 2820–2830. doi:10.1109/TITS.2018.2873092.

Li, X., Hu, S., Fan, W., Deng, K., 2018b. Modeling an enhanced ridesharing system with meet points and time windows. PloS one 13, e0195927. doi:10.1371/journal.pone.0195927.

Liang, X., de Almeida Correia, G.H., An, K., van Arem, B., 2020. Automated taxis' dial-a-ride problem with ride-sharing considering congestion-based dynamic travel times. Transportation Research Part C: Emerging Technologies 112, 260–281. doi:10.1016/j.trc.2020.01.024.

Lyu, Y., Lee, V.C., Ng, J.K.Y., Lim, B.Y., Liu, K., Chen, C., 2019. Flexi-sharing: a flexible and personalized taxi-sharing system. IEEE Transactions on Vehicular Technology 68, 9399–9413. doi:10.1109/TVT.2019.2932869.

Ma, S., Zheng, Y., Wolfson, O., 2013. T-share: A large-scale dynamic taxi ridesharing service, in: 2013 IEEE 29th International Conference on Data Engineering (ICDE), IEEE. pp. 410–421. doi:10.1109/ICDE.2013.6544843.

Manjunath, A., Raychoudhury, V., Saha, S., Kar, S., Kamath, A., 2021. Care-share: a cooperative and adaptive strategy for distributed taxi ride sharing. IEEE Transactions on Intelligent Transportation Systems doi:10.1109/TITS.2021.3066439.

Martínez, L.M., Viegas, J.M., Eiró, T., 2015. Formulating a new express minibus service design problem as a clustering problem. Transportation Science 49, 85–98.

Miklas-Kalczynska, M., Kalczynski, P., 2021. Self-organized carpools with meeting points. International Journal of Sustainable Transportation 15, 140–151.

Moreira-Matias, L., Gama, J., Ferreira, M., Damas, L., 2012. A predictive model for the passenger demand on a taxi network, in: 2012 15th International IEEE Conference on Intelligent Transportation Systems, IEEE. pp. 1014–1019.

Moreira-Matias, L., Gama, J., Ferreira, M., Moreira, J., Damas, L., 2013. Predicting taxi-passenger demand using streaming data. IEEE Transactions on Intelligent Transportation Systems 14, 1393–1402. doi:10.1109/TITS.2013.2262376.

Mourad, A., Puchinger, J., Chu, C., 2019. A survey of models and algorithms for optimizing shared mobility. Transportation Research Part B: Methodological 123, 323–346. doi:10.1016/j.trb.2019.02.003.

NYC City Taxi & Limousine Commission: TLC, 2014. 2014 taxicab fact book. URL: https://www.nyc.gov/assets/tlc/downloads/pdf/2014_tlc_factbook.pdf. accessed: 07.09.2023.

NYC City Taxi & Limousine Commission: TLC, 2017. Tlc trip record data. URL: https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page. accessed: 14.01.2022.

NYC City Taxi & Limousine Commission: TLC, 2023. Tlc cost components. URL: https://www.nyc.gov/site/tlc/passengers/taxi-fare.page. accessed: 28.07.2023.

Qian, X., Zhang, W., Ukkusuri, S.V., Yang, C., 2017. Optimal assignment and incentive design in the taxi group ride problem. Transportation Research Part B: Methodological 103, 208–226. doi:10.1016/j.trb.2017.03.001.

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.

Santos, D.O., Xavier, E.C., 2015. Taxi and ride sharing: A dynamic dial-a-ride problem with money as an incentive. Expert Systems with Applications 42, 6728–6737. doi:10.1016/j.eswa.2015.04.060.

Silwal, S., Gani, M.O., Raychoudhury, V., 2019. A survey of taxi ride sharing system architectures, in: 2019 IEEE International Conference on Smart Computing (SMARTCOMP), IEEE. pp. 144–149. doi:10.1109/SMARTCOMP.2019.00044.

Smet, P., 2021. Ride sharing with flexible participants: a metaheuristic approach for large-scale problems. International Transactions in Operational Research 28, 91–118. doi:10.1111/itor.12737.

Stiglic, M., Agatz, N., Savelsbergh, M., Gradisar, M., 2015. The benefits of meeting points in ride-sharing systems. Transportation Research Part B: Methodological 82, 36–53. doi:10.1016/j.trb.2015.07.025.

Stiglic, M., Agatz, N., Savelsbergh, M., Gradisar, M., 2016. Making dynamic ride-sharing work: The impact of driver and rider flexibility. Transportation

Research Part E: Logistics and Transportation Review 91, 190–207. doi:10.1016/j.tre.2016.04.010.

Stiglic, M., Agatz, N., Savelsbergh, M., Gradisar, M., 2018. Enhancing urban mobility: Integrating ride-sharing and public transit. Computers & Operations Research 90, 12–21. doi:10.1016/j.cor.2017.08.016.

Stock, E., 2018. Introducing express pool: Walk a little to save a lot. https://www.uber.com/newsroom/expresspool/. Accessed 07-February-2023.

Sun, Y., Chen, Z.L., Zhang, L., 2020. Nonprofit peer-to-peer ridesharing optimization. Transportation Research Part E: Logistics and Transportation Review 142, 102053. doi:10.1016/j.tre.2020.102053.

Tafreshian, A., Masoud, N., Yin, Y., 2020. Frontiers in service science: Ride matching for peer-to-peer ride sharing: A review and future directions. Service Science 12, 44–60. doi:10.1287/serv.2020.0258.

Ting, K.H., Lee, L.S., Pickl, S., Seow, H.V., 2021. Shared mobility problems: A systematic review on types, variants, characteristics, and solution approaches. Applied Sciences 11, 7996. doi:10.3390/app11177996.

Wang, S., Correia, G.H.d.A., Lin, H.X., 2019. Exploring the performance of different on-demand transit services provided by a fleet of shared automated vehicles: An agent-based model. Journal of Advanced Transportation 2019. doi:10.1155/2019/7878042.

Wang, X., Agatz, N., Erera, A., 2018. Stable matching for dynamic ride-sharing systems. Transportation Science 52, 850–867. doi:10.1287/trsc.2017.0768.

Wang, Y., Li, M., 2021. Optimization algorithm design for the taxi-sharing problem and application. Mathematical Problems in Engineering 2021. doi:10.1155/2021/5572200.

Wang, Z., Hyland, M.F., Bahk, Y., Sarma, N.J., 2022. On optimizing shared-ride mobility services with walking legs. arXiv preprint arXiv:2201.12639 .

Weber, J., 2020. Bewegende Zeiten: Mobilität der Zukunft. Springer. doi:10.1007/978-3-658-30311-2.

Yan, C., Zhu, H., Korolko, N., Woodard, D., 2020. Dynamic pricing and matching in ride-hailing platforms. Naval Research Logistics (NRL) 67, 705–724.

Yuan, J., Zheng, Y., Zhang, L., Xie, X., Sun, G., 2011. Where to find my next passenger, in: Proceedings of the 13th international conference on Ubiquitous computing, pp. 109–118. doi:10.1145/2030112.2030128.

Zhao, M., Yin, J., An, S., Wang, J., Feng, D., 2018. Ridesharing problem with flexible pickup and delivery locations for app-based transportation service: Mathematical modeling and decomposition methods. Journal of Advanced Transportation 2018. doi:10.1155/2018/6430950.

Appendix A. Results

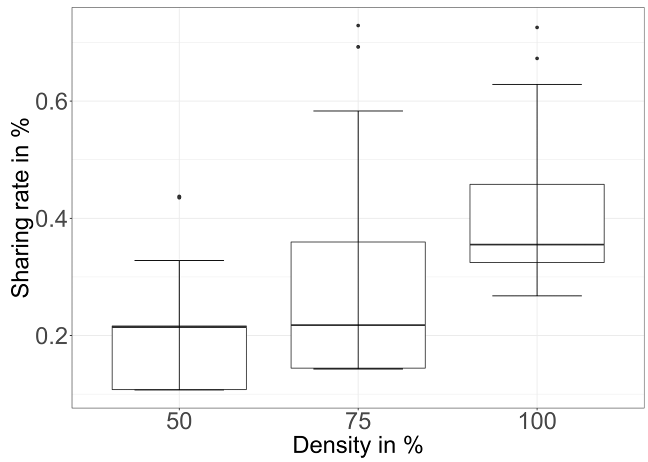

Figure A.7: Boxplot showing sharing rate (in %) in Porto

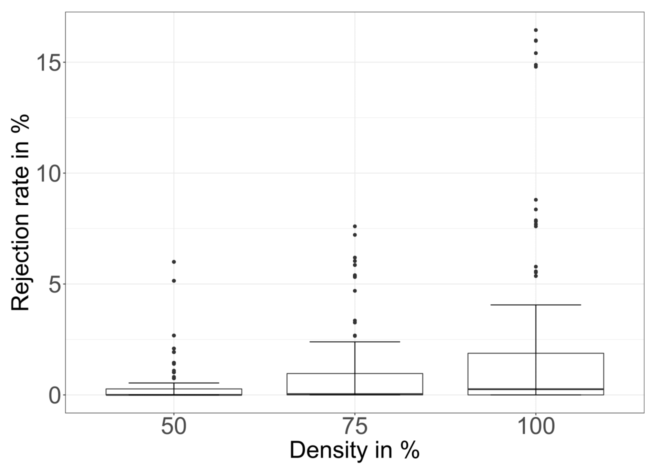

Figure A.8: Boxplot showing rejection rate (in %) in Porto

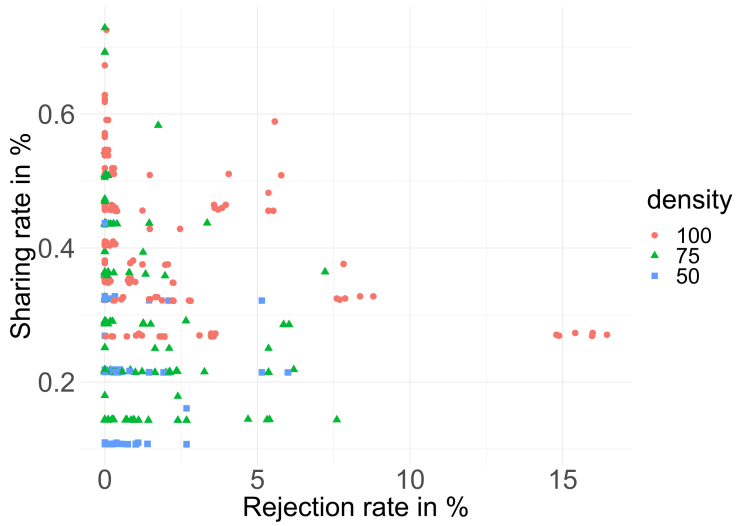

Figure A.9: Relation between sharing and rejection rate in Porto

Table A.8: Gap Statistics

	NYC	Porto
Mean	0.000	0.000
Median	0.000	0.000
Min	0.000	0.000
Max	0.011	0.000
Stdv	0.000	0.000