WORKING PAPER SERIES

Heatmap-based Decision Support for Repositioning in Ride-Sharing Systems

Jarmo Haferkamp/Marlin Wolf Ulmer/Jan Fabian Ehmke

Working Paper No. 3/2022

OTTO VON GUERICKE
UNIVERSITÄT
MAGDEBURG

FACULTY OF ECONOMICS
AND MANAGEMENT

Impressum (§ 5 TMG)

Herausgeber:
Otto-von-Guericke-Universität Magdeburg
Fakultät für Wirtschaftswissenschaft
Der Dekan

Verantwortlich für diese Ausgabe:
J.Haferkamp, M.W.Ulmer and Jan F. Ehmke
Otto-von-Guericke-Universität Magdeburg
Fakultät für Wirtschaftswissenschaft
Postfach 4120
39016 Magdeburg
Germany

http://www.fww.ovgu.de/femm

Bezug über den Herausgeber
ISSN 1615-4274

Heatmap-based Decision Support for Repositioning in Ride-Sharing Systems

Jarmo Haferkamp
Otto von Guericke University Magdeburg, 39106 Magdeburg, Germany

Marlin Wolf Ulmer
Otto von Guericke University Magdeburg, 39106 Magdeburg, Germany

Jan Fabian Ehmke
University of Vienna, Business Decisions and Analytics, 1090 Vienna, Austria
University of Vienna, Research Network Data Science, 1090 Vienna, Austria

In ride-sharing systems, platform providers aim to distribute the drivers in the city to meet current and potential future demand and to avoid service cancellations. Ensuring such distribution is particularly challenging in the case of a crowdsourced fleet, as drivers are not centrally controlled but are free to decide where to reposition when idle. Thus, providers look for alternative ways to ensure a vehicle distribution that benefits both users and drivers, and consequently the provider. We propose an intuitive means to improve idle ride-sharing vehicles' repositioning: repositioning opportunity heatmaps. These heatmaps highlight driver-specific earning opportunities approximated based on the expected future demand, fleet distribution, and location of the specific driver. Based on the heatmaps, drivers make decentralized yet better-informed repositioning decisions. As our heatmap policy changes the driver distribution, we propose an adaptive learning algorithm for designing our heatmaps in large-scale ride-sharing systems. We simulate the system and generate heatmaps based on previously learned repositioning opportunities in every iteration. We then update these based on the simulation's outcome and use the updated values in the next iteration. We test our heatmap design in a comprehensive case study on New York ride-sharing data. We show that carefully designed heatmaps reduce service cancellations therefore revenue loss for platform and drivers significantly.

Key words: mobility-on-demand, vehicle repositioning, crowdsourced transportation, heatmap, stochastic dynamic decision making, adaptive learning

1. Introduction

The trend of ride-sharing is unbroken. Services like UberXShare (*https://www.uber.com/us/en/ride/uberpool*) and Berlkönig (*https://www.berlkoenig.de*) offer convenient and affordable services at low costs. In these systems, ride-sharing users spontaneously submit transportation requests online, are picked up a short time afterward, and driven to their destination, while possibly sharing part of their ride with other users. In some cases, it may not be possible to provide a sufficient level of service because no driver from the ride-sharing fleet is

currently available in the vicinity of the user, creating high waiting times. Service cancellations due to insufficient levels of service lead to loss of revenue and customer dissatisfaction. Thus, service providers aim for a good distribution of the drivers in the city to meet current and potential future demand.

The tools for ensuring such distribution depend on the type of service provider. Services associated with public transport like Berlkönig as well as some private providers own a fleet of employed drivers and, therefore, can make repositioning decisions centrally. Since the future demand is uncertain, efficient repositioning decisions are a challenging task on their own, even if they can be made centrally (Pouls, Meyer, and Ahuja 2020). Other providers such as UberXShare crowdsource transportation for their ride-sharing services to private individuals that are paid on a per-job basis. In that case, drivers are not controlled by the service provider directly but are free to decide where to reposition in the city when unoccupied. This is an additional challenge, as decentralized repositioning likely leads to inconvenience for users and drivers. For example, many drivers may prefer waiting in the city center where new requests are more likely to occur, which may lead to an overflow of resources, while there is a driver shortage at other locations. This means poor service availability for users in other areas of the city, and fewer earning opportunities for the many drivers in the city center. Another challenge with crowdsourced drivers is that they are reluctant to follow directions by the platform if they do not understand them (Möhlmann et al. 2020). Thus, service providers look for alternative and intuitive ways to ensure a good distribution of crowdsourced drivers that benefits users, drivers, and the provider.

In this paper, we propose an intuitive means to improve the repositioning of unoccupied vehicles in ride-sharing systems: Repositioning Opportunity Heatmaps (ROH). These heatmaps highlight driver-specific earning opportunities, approximated based on the travel time to areas with a shortage of drivers to satisfy the expected demand in the near future. In the heatmaps, repositioning locations with high expected opportunities are shown in green shades while repositioning locations with low expectations are shown in red shades. Both shades can have different intensities, e.g., dependent on the relative opportunity volumes. Based on the heatmaps, drivers then make decisions on repositioning in a decentral and independent manner. Ideally, heatmaps guide drivers intuitively to increase service availability, reduce cancellations, and improve drivers' earning opportunities.

Creating heatmaps is challenging for several reasons. As with centralized approaches, future user demand is uncertain, therefore, the earning opportunities are unknown. Furthermore, showing heatmaps changes the repositioning decisions of drivers in the system, which in turn may lead again to too many or not enough drivers in certain areas. To address this issue, we propose an adaptive learning algorithm for designing our heatmaps. In every iteration, we simulate the system and

generate heatmaps based on previously learned opportunities. We then update the earning opportunities based on the simulation's outcome and use the updated opportunities in the next iteration. Eventually, the expected opportunities and therefore the heatmap design policy converges.

We test our heatmap design in a comprehensive case study on New York ride-sharing data with 200 drivers and around 6400 expected users per planning period. We show that carefully designed heatmaps reduce service cancellations therefore revenue loss for platform and drivers significantly. Furthermore, providing heatmaps to drivers does not only increase the average earnings per driver but also reduces the volatility in earnings among the drivers. Even though an analysis of these results is beyond the scope of this paper, such "fair" earnings might be an important factor for the longer-term commitment of a driver to a platform. We also show that heatmaps not only lead to a better but also more balanced distribution of service availability in the city, another important factor for long-term user retention. Finally, we analyze how the adoption or denial of the heatmap's recommendations impact driver earnings locally and globally. The overall system performs best when all drivers are relatively compliant, but results show that there is potential for and danger of single drivers gaming the system. Our paper makes the following contributions:

- We investigate and define a new problem and formalize the corresponding sequential decision process.
- We are among the first to introduce heatmaps for nudging crowdsourced drivers. We show that heatmaps are a powerful, intuitive tool for the management of complex and dynamic systems, which is likely applicable for other applications such as restaurant meal delivery.
- We present a large-scale, real-time decision policy based on adaptive learning with very limited calculation efforts.
- We show that our policy is superior to a variety of benchmark policies and partially outperforms even state-of-the-art centralized repositioning policies.
- We provide a comprehensive computational study including valuable managerial insights on central and decentral management of ride-sharing systems.

We begin with an overview of related literature in Section 2. The repositioning problem is stated and formalized in Section 3. How we develop repositioning heatmaps by adaptive learning is presented in Section 4. The experimental setup with modeling of compliant and non-compliant driver behavior is discussed in Section 5. Computational experiments are reported in Section 6. We conclude with final remarks in Section 7.

2. Related Literature

Our work addresses the management of crowdsourced drivers for ride-sharing services. In the following, we will briefly discuss the (very limited) related work from crowdsourced transportation, and then embed our work in the literature on ride-sharing optimization.

2.1. Crowdsourced Transportation

Crowdsourced transportation service providers outsource jobs to private individuals, inducing cost advantages for the service provider on the one hand, and flexible working hours and uncertain earning opportunities for the drivers on the other hand. Uncertainty in crowdsourced transportation plays a major role for service providers as well, as it is not clear when, where, and how many drivers are available. Moreover, the freedom of choice of the drivers leads to further planning uncertainty, since they decide on the acceptance of a job assignment and the execution of a repositioning recommendation. Optimization in the face of these uncertainties is a major challenge that has only recently come into the focus of transportation research. For a recent overview, we refer to Savelsbergh and Ulmer (2022). The rather limited work in this area focuses on the uncertain number of drivers being in the system (see, e.g., Dayarian and Savelsbergh 2020 or Ulmer and Savelsbergh 2020) or hedging against drivers rejecting offered requests (see, e.g., Gdowska, Viana, and Pedroso 2018 or Ausseil, Pazour, and Ulmer 2022).

Uncertainty in repositioning has not been explored much so far. One of the few works that address repositioning under these conditions is Alnaggar (2021). Similar to our work Alnaggar (2021) proposes using a heatmap to guide drivers, in their case, for crowdsourced last-mile deliveries. They use a short-term demand forecast to derive global heatmaps with up to three different levels, for systems with up to nine repositioning locations. We differ from their work as follows. First, we consider the problem of ride-sharing with consolidation potential and tighter time commitments. Second, our work differs in scale with hundreds of repositioning locations and drivers and thousands of customers. Third, we propose heatmaps of continuous granularity designed for individual drivers.

2.2. Ride-Sharing

In the following, we present related work from ride-sharing, which suffers from an unbalanced driver distribution due to spatial and temporal imbalanced demand (Jiao et al. 2021). Besides the assignment of transport requests, imbalanced demand is one of the most considered challenges for the efficient operation of ride-sharing services (Wang, Shim, and Wu 2019). In the following, we examine the extent to which the literature for large-scale repositioning of unoccupied drivers in ride-sharing services meets the requirements of a crowdsourced fleet (see Table 1) .

With regard to the unique requirements of crowdsourced fleets, four criteria have been identified as particularly relevant: (1) Pursuing a system-wide balance between demand and supply in the best interest of provider revenue and driver earning opportunities; (2) The balance is achieved by supporting drivers in their decision about when and where to reposition; (3) The recommendations include comprehensive information to enable well-founded decision making (no take it or leave it); (4) Non-compliance is not penalized to ensure long-term driver satisfaction and therefore retention.

		Platform-wide demand-supply balance	Supports drivers in decision making	Indicates multiple repositioning options	No discrimination against inconvenient drivers/users
Pavone et al. (2012) Zhang and Pavone (2016) Sayarshad and Chow (2017) Braverman et al. (2019)	Queueing-based	√			√
Zhang, Rossi, and Pavone (2016) Iglesias et al. (2018) Wallar et al. (2018) Pouls, Meyer, and Ahuja (2020) Lei, Qian, and Ukkusuri (2020) Li et al. (2021)	Model predictive control	√			√
Wen, Zhao, and Jaillet (2017) Holler et al. (2019) Jiao et al. (2021) Liu, Chen, and Chen (2021) Xi et al. (2021) Yu and Hu (2021)	Reinforcement learning	√	√		√
Li et al. (2011) Powell et al. (2011) Yuan et al. (2011)	Data-based taxi guidance		√	√	√
Bimpikis, Candogan, and Saban (2019) Guda and Subramanian (2019) Hu, Hu, and Zhu (2021)	Surge pricing	√	√	√	
ROH	Heatmaps	√	√	√	√

Table 1 Literature classification

The related approaches can be divided into five main streams: queueing-based, model predictive control, reinforcement learning, data-based driver guidance, and surge pricing. In the following, we briefly present the related work per category and highlight the relationship to our work.

We first consider *queueing-based approaches* and *model predictive control*, as they are similar in their applicability for crowdsourced fleets. Queueing-based approaches for determining optimal repositioning policies have been widely researched (e.g. Pavone et al. (2012), Zhang and Pavone (2016), Sayarshad and Chow (2017), and Braverman et al. (2019)). Sayarshad and Chow (2017) proved that these are also applicable for services of real-world size. Braverman et al. (2019) show that, in addition to classical repositioning, the optimal routing for unoccupied vehicles in search of a next passenger can be determined. Model predictive control typically builds on demand forecasts combined with mathematical programming to solve the repositioning problem online periodically. The first contribution in this direction comes from Zhang, Rossi, and Pavone (2016). Others like Iglesias et al. (2018), Wallar et al. (2018), Pouls, Meyer, and Ahuja (2020) focus on large-scale ride-sharing or shared rides. Lei, Qian, and Ukkusuri (2020) propose model predictive control to train a neural network offline that allows for a quick repositioning policy prediction online. Li et al. (2021) also propose a neural network but use it to improve the demand prediction within a model predictive control approach. Although queueing-based approaches and model predictive control are based on very different concepts of ride-sharing systems, they share the general assumption that the system is controlled centrally and that drivers play no or only a minor role in repositioning decisions. They are hence inapplicable for guiding individual drivers in a crowdsourced ride-sharing system. In contrast, with the proposed heatmap, we offer a decision support tool at the request of a driver that helps to make better-informed decisions on an individual basis.

A different approach for repositioning in large-scale ride-sharing services that gained attention in recent years is *agent-based reinforcement learning* (e.g. Wen, Zhao, and Jaillet (2017), Holler et al. (2019), and Jiao et al. (2021)). In agent-based approaches, decentralized decision making for individual drivers is considered instead of centralized decisions for the entire fleet. While Wen, Zhao, and Jaillet (2017) focus on decentralized learning of a repositioning policy, Holler et al. (2019) compare the benefits of both centralized and decentralized learning. Of particular interest for our work are the results of Jiao et al. (2021). The results show that policies learned via reinforcement learning can be transferred to the real world and are advantageous over the intuitive decisions of drivers. It is further shown that driver collaboration is particularly relevant for large fleets. Moreover, it is argued that driver repositioning may deviate from central recommendations, but this has not been considered further in the experiments. Finally, recent contributions by Liu, Chen, and Chen (2021), Xi et al. (2021), and Yu and Hu (2021) propose technical innovations for reinforcement learning of repositioning policies.

As Jiao et al. (2021) have shown, reinforcement learning can be a good foundation for learning recommendations for the repositioning of drivers. However, due to the nature of reinforcement learning, the recommendations as described in the above papers are limited to the driving direction to be taken or the neighboring area to be approached. Furthermore, these are overall black-box approaches. Drivers therefore only have the choice of complying with recommendations or making completely unsupported repositioning decisions. Our repositioning heatmaps differ from this, despite their learning capability, by providing drivers with an assessment of repositioning locations distributed across the service area while highlighting the best individual option. Moreover, the information used to generate the heatmap as well as the resulting indications can be communicated in a comprehensible way.

Further papers have focused on providing taxi guidance by indicating the *profitability* of different repositioning options (e.g., Li et al. (2011), Powell et al. (2011), and Yuan et al. (2011)). These papers are based on the analysis of large amounts of historical booking data in order to derive profitable behavioral strategies. They have in common that a profitability maximizing behavior is considered only for one driver without reflecting the distribution of the fleet. Demand-supply balancing to maximize platform revenue is not considered; instead, having all drivers follow the recommended strategy may lead to larger imbalances and less total revenue. We, therefore, integrate a learning process as well as fleet-related information that ensures balancing recommendations by the heatmap from which both service providers and all drivers can benefit.

Finally, another way to rebalance a crowdsourced fleet while focusing on platforms' revenue is *surge pricing*, where user fees (and therefore driver compensation) are increased to attract drivers (e.g. Bimpikis, Candogan, and Saban (2019), Guda and Subramanian (2019), and Hu, Hu, and

Zhu (2021)). In this context, heatmaps have already been applied to indicate drivers of the current distribution of prices (Guda and Subramanian 2019). This provides drivers with a strong monetary incentive to behave as the platform requires. However, surge pricing is usually used to serve peak demand and does not necessarily increase service availability in every part of the city (Bimpikis, Candogan, and Saban 2019). Furthermore, volatile prices are often not well received by users and drivers (see e.g. Dholakia (2015), Goncharova (2017), and Conger (2021)). As a result, first cities have begun to ban surge pricing (Spielman 2021). With our repositioning heatmap, we, therefore, aim to provide a less controversial means supporting ride-sharing service providers in managing a crowdsourced fleet without inducing disadvantages for particular driver or user groups.

3. Problem Statement and Formulation

In the following, we give a problem narrative to define the system and its dynamics. We then provide an example to illustrate the sequential decision process model.

3.1. Problem Narrative

We take the perspective of a ride-sharing service provider that connects transportation requests with self-employed drivers following the goal of minimizing the overall number of daily service cancellations. Over the course of the service horizon, users request instant transportation from an origin to a destination within the city. To fulfill a transportation request, the service provider can assign a nearby driver that is either currently idling, or assign an already busy driver who is currently transporting another user. Thus, pooling is possible with respect to a maximum vehicle capacity. If no driver is nearby, e.g., the user cannot be picked up within 10 minutes, the user cancels the service request. If the transportation request can be fulfilled, the user pays a fee. This fee is split between driver and service provider. For simplicity, we assume that the fleet of self-employed drivers work the entire service horizon, accept all transportation requests assigned to them, and follow the routing suggested by the service provider. However, the drivers are free in their decision where to reposition after finishing a job, which could be an area with high expected demand (as described by Ermagun and Stathopoulos (2018), for example), or an area in the driver's neighborhood (as described by Rai, Verlinde, and Macharis (2021), for example).

We assume decentralized decision-making about the repositioning of the drivers. Thus, the drivers' repositioning behavior impacts the fleet distribution in the city and therefore future revenue respectively earning opportunities for service providers and drivers. Having not enough drivers in one area of the city leads to service cancellations and revenue loss for the service provider. Having too many drivers in another area leads to fewer earning opportunities for the drivers. Therefore, both service providers and drivers have some interest in the effective distribution of the fleet. However, research shows that self-employed drivers are reluctant to follow service providers' directions

straightforwardly, especially if they seem counterintuitive and their reasoning is not immediately clear (Möhlmann et al. 2020). Hence, in our problem, we propose an alternative tool that is both intuitive and leaves drivers with the final repositioning decision: driver guidance through heatmaps.

Driver-specific heatmaps indicate earning opportunities for the drivers in the city. Earning opportunities depend on the repositioning time from the driver's current to the repositioning location, the expected demand in the area as well as how well the area is already covered by the fleet. Promising repositioning locations imply significant opportunities due to short travel times, high expected demand, and/or few drivers in the surrounding area. These repositioning locations are indicated in shades of green with different intensities; other, rather unfavorable repositioning locations are colored in different shades of red. Whenever drivers are without an assignment to fulfill and consider repositioning, they consult their heatmap to make a well-informed repositioning decision.

The service provider searches for a heatmap strategy that nudges drivers towards lucrative areas and still maintains a flexible and effective distribution of the entire fleet to avoid future service cancellations. Such a strategy creates a heatmap every time a driver becomes idle, based on the current state information and the expected demand. Since our work focuses on heatmap design, we assume that for assignment and routing, the service provider follows an externally given strategy (the optimization of this strategy is outside of the scope of this paper).

3.2. Example

In the following, we give an example to illustrate the dynamics of the system under consideration (see Figure 1). The system can be described by its state (on the left), potential decisions (in the center), and a realization of stochastic information including the resulting transition to a new state (on the right). The example is at time $t = 60$ of the service horizon. The city consists of four regions, each with one repositioning location. The locations of the three drivers are given. Two of them, drivers two and three, depicted in light, already have done their repositioning and are now waiting for new requests to be assigned. Driver one, depicted in dark, just finished a trip and now needs to make a repositioning decision via a heatmap provided by the service provider.

A potential heatmap recommendation is shown in the center. As the two areas on the right are already covered by a driver, a reasonable decision would be to color these locations in red to indicate fewer earning opportunities. The areas on the left are both uncovered currently, thus, the heatmap highlights them in green, indicating more earning opportunities for these two locations. We note that in the real problem and in the following model and method, the levels of color intensities are continuous and therefore, the number of decisions per state is infinite.

The final part of Figure 1 shows the realization of stochastic information and the transition to the next state at time $t = 75$. This part is relatively complex for this problem. The stochastic information is twofold:

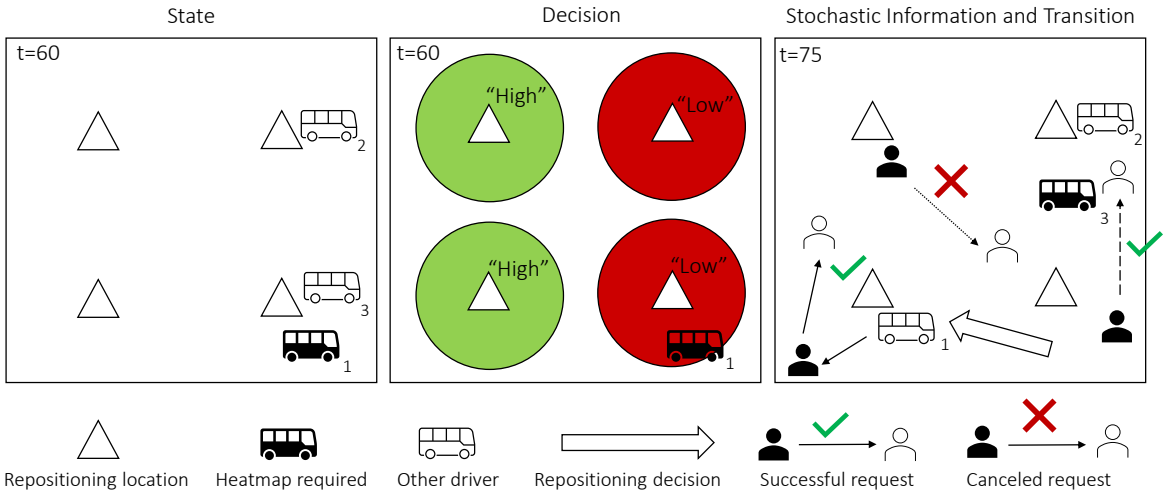


Figure 1 **Example for a state, decision, the realization of stochastic information, and transition to the next state**

1. First, the repositioning decision of driver one has been revealed, indicated by the large white arrow. We assume that the driver informs the service provider about the selected repositioning location as it is beneficial for all parties. The platform can provide a more informative heatmap and as a result, the driver has less competition to worry about. Here, driver one has decided to reposition to the bottom left, likely because expected earning opportunities are higher than in the right corner and the trip was shorter than to the upper left corner.
2. Second, new demand occurs (and is assigned to drivers if possible) between the current and new decision epoch, indicated by dark and light customer icons connected by a thin arrow. In the example, three new requests occurred between $t = 60$ and $t = 75$. Two of them could be assigned. The one in the bottom right was served by driver three, the other in the bottom left was served by driver one. The request in the center of the city could not be assigned and was canceled. As the platform aims on minimizing cancellations, the cost realized in the transition is one.

The new decision epoch is initialized when the next driver, namely driver three in this example, requires a repositioning heatmap. As the demand is uncertain, it is also uncertain when the next decision epoch occurs ($t = 75$ in the example). The transition leads to a new distribution of the fleet, based on the repositioning decision of driver one and the assignments of users to drivers and the corresponding trips.

3.3. Sequential decision process

In the following, we present the sequential decision process for the problem, i.e., the definition of decision epochs, states, decisions, stochastic information, and transitions. First, we introduce some preliminary notations.

Preliminaries. We assume operations during a time horizon $T = [0, t^{\max}]$, where time units are discretized by rounding down to minutes. Operations take place in a service area $(N, E, \mathcal{T})$ with N being the set of locations in the city, E the set of edges between the locations, and $\mathcal{T}$ the constant travel times on the edges. We define a set of potential repositioning locations $R \subset N$. We further assume a fleet of m drivers working all day with initial idling positions $\rho_0 = (\rho_{01}, \ldots, \rho_{0m}) \in R^m$ and homogeneous capacity and service duration per stop. There is no termination location, but all operations end (latest) at time $t^{\max}$.

Decision epochs. Decision epochs occur whenever a driver finished service and checks the app for a heatmap. Thus, the time of the next decision epoch and the overall number of decision epochs is uncertain. We denote decision epochs as $k = 1, \ldots, K$ with K being a random variable.

States. A state S_k at decision point k contains the following information:

- The current point of time t_k.
- The location information of the driver currently requesting a heatmap: n_k.
- The status of the other drivers: Given the large scale of the underlying problem in drivers and customers, we refrain from modeling the individual routes and stops for every driver (and customer) in the system. Instead, we focus on the u_k currently unoccupied drivers either idling at or traveling to a repositioning location. We model the corresponding information using vectors $\rho_k = (r_{1k}, \ldots, \rho_{u_k k})$ and $\tau_k = (\tau_{1k}, \ldots, \tau_{u_k k})$ to represent locations and time. For driver i, $\rho_{ik} \in R$ indicates the repositioning location and τ_{ik} indicates the time the driver will arrive there. Notably, in case of $\tau_{ik} = t_k$, the driver is already waiting at the location ρ_{ik} while in case of $\tau_{ik} > t_k$, the driver is still on the way.

In essence, a state can be summarized as $S_k = (t_k, n_k, \rho_k, \tau_k)$. The initial state is at time zero with no driver requesting a heatmap and all drivers idling at their initial locations, $S_0 = (0, -, \rho_0, \vec{0})$. There is no decision made in the initial state.

Decisions. A decision x_k is the heatmap of repositioning opportunities shown to the driver. A heatmap decision is a vector $x_k = (x_{1k}, \ldots, x_{|R|k})$ of values $x_{rk} \in \mathbb{R}_+$ for each repositioning location $r \in R$. Higher values of x_{rk} indicate higher opportunities. There are no direct costs associated with a decision.

Stochastic Information and Transition. The stochastic information $w_{k+1} = (r^w_{k+1}, D_{k+1})$ is twofold and reflects repositioning and the occurrence and treatment of new demand:

First, it reveals a new repositioning location r^w_{k+1} for the driver requesting a heatmap in S_k based on heatmap decision x_k. In our experiments, we assume that the probability of location $r \in R$ being selected is based on the value x_{rk} (higher probability with higher value).

Second, demand D_{k+1} realized and is served via the platform's assignment and routing procedure until another driver requests a heatmap. The demand is realized until a point of time t_{k+1} where the assignment and routing procedure induces the next free driver requesting a heatmap and therefore the next decision epoch $k+1$ at time t_{k+1}. Furthermore, cancellations D^c_{k+1} between t_k and t_{k+1} due to insufficient driver availability realize with information about time t^p and the nearest repositioning location $l^p \in R$ for a cancellation p. The cancellations define the cost $C(S_k, x_k, D_{k+1}) = |D^c_{k+1}|$.

In case $t_{k+1} = t^{\max}$, the process terminates.

Solution and Objective Function. The solution for the problem is a policy π assigning a heatmap decision $\mathcal{X}^\pi(S_k)$ to every state S_k. An optimal policy π^* minimizes the expected costs (cancellations) when starting in state S_0 and applying policy π^* throughout the process:

$$\pi^* = \underset{\pi \in \Pi}{\arg\min}\, \mathbb{E}\left[\sum_{k=0}^{K} C(S_k, \mathcal{X}^\pi(S_k)) | S_0\right]. \tag{1}$$

4. Repositioning Opportunity Heatmaps

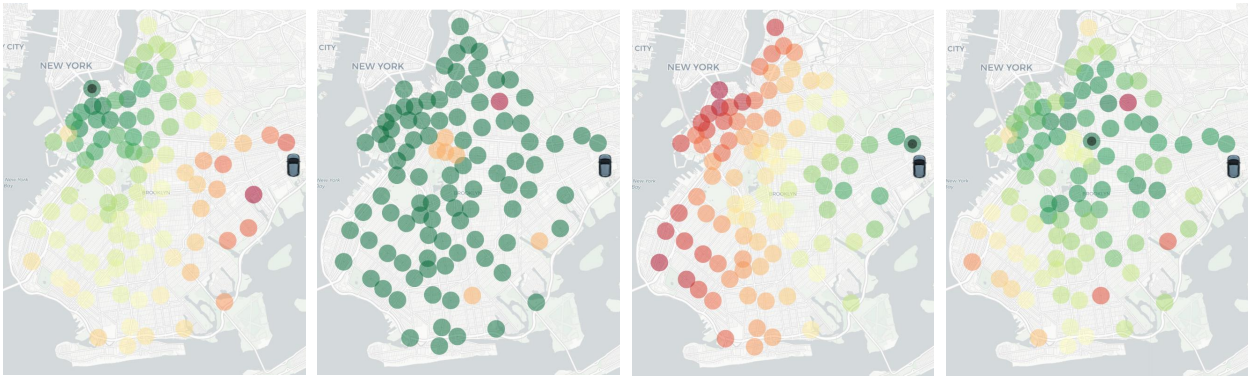
In the following, we present the methodology behind creating repositioning opportunity heatmaps. We first give a general motivation for and overview of our method and then present the details.

4.1. Motivation and Overview

Designing heatmaps is challenging for several reasons. The future earning opportunities at repositioning location r for the driver requesting a heatmap in k depend on three factors:

- The *expected cancellations* that drivers can reduce when being repositioned to r in the "near" future. Higher cancellations will increase earning opportunities. However, expected cancellations in the far-away future may be less relevant and may lead to unnecessary waiting for the driver as well as missed earning opportunities and future cancellations elsewhere.
- The *unoccupied drivers* that are currently idling at or that are on their way to location r. A larger number will decrease earning opportunities for the drivers. At the same time, sending another driver will likely lead to driver shortage and future cancellations in other areas of the city.
- The *travel time* for the driver located n_k to r. A longer trip will decrease earning opportunities and increase future cancellations as the driver is occupied traveling.

Consequently, we design heatmaps based on the expected future cancellations, the current distribution of drivers at relocation locations, and the position of the driver requesting the heatmap (We will later show that each part is essential for an effective heatmap design). Integrating all these intuitive factors in one holistic heatmap is already difficult as we discuss in the following section. However, we face an additional challenge, namely, that the expected cancellations (and consequently, the future earning opportunities) are also affected by the policy applied since the



(a) Expected Cancellations (b) Distribution of Idle Drivers (c) Repositioning Times (d) Repositioning Opportunity Heatmap

Figure 2 Exemplary Repositioning Heatmap

policy impacts not only the repositioning location of the driver in the current state but also the distribution of all drivers due to the heatmaps shown to them in later states.

Integrating these large-scale stochastic dynamic developments in the heatmap design is difficult, especially, since the intuitive composition of the heatmap should be maintained. To this end, we propose an adaptive learning procedure that does not change the design process of the heatmap itself but carefully adapts its most important component, the expected cancellations. The learning procedure starts with initial expected cancellations values, iteratively applies the corresponding policy, and adapts the expected cancellations based on the observed values. Therefore, the resulting heatmaps maintain their intuition but integrate the stochastic dynamic developments at the same time. The details of the learning process are presented later in this section, after introducing the initial holistic heatmap design.

4.2. Heatmap Design

The intention of our heatmap design is to link state and learned information in a smart way to provide superior and intuitive repositioning recommendations. The resulting advantages are better explainability, adaptability, and reproducibility of our heatmaps. To achieve this, the repositioning heatmap is designed as a combination of three underlying heatmaps: one reflecting information on the expected cancellations, one reflecting the fleet distribution, and one reflecting the travel times for the driver under consideration. Before formalizing this process, we illustrate how a heatmap is created (see Figure 2):

2(a) The first heatmap shows the distribution of the learned expected cancellations given the current point of time. The heatmap thus indicates expected future demand in the neighborhood of repositioning locations as well as the demand that is covered by future driver movements. Consequently, it combines both pieces of information to identify opportunities for proactive repositioning. In the example, in the northwest high expected cancellations and therefore

opportunities are highlighted in green, while in the eastern regions where the requesting driver is located red colors indicate few opportunities for the near future. However, the black-filled point in the far north shows that this heatmap alone would recommend a very time-consuming repositioning.

2(b) The second heatmap is state-dependent. It displays the current distribution of the idle fleet among the repositioning locations. It, therefore, indicates how many competing drivers are to be expected and thus the extent to which expected opportunities are already covered. In the example heatmap, unoccupied repositioning locations are colored in green, and the most frequented ones are in red. Since the majority of repositioning locations are unoccupied, no recommendation can be derived from this heatmap alone.

2(c) The third heatmap indicates the travel time between the driver's current location and the potential repositioning locations. Including travel times in the heatmap design personalizes repositioning opportunities and counterbalances the expected cancellations. Accordingly, the repositioning locations are colored in the example from green to red with increasing distance. Based on this heatmap, the nearest repositioning locations would always be recommended.

2(d) Based on these three pieces of information visualized by heatmaps, the final repositioning heatmap shown in Figure 2(d) is created. As can be noticed, a well-balanced recommendation results from the three one-dimensional heatmaps, considering a trade-off between expected cancellations and travel time while avoiding direct competitors.

After understanding how the heatmap should be constructed practically, we formalize the generation in the following. Let us first consider again the expected cancellation heatmap shown in Figure 2(a). The information displayed here is based on a learned matrix $c = (c_{rt})_{r \in R, t \in T}$ of values $c_{rt} \in \mathbb{R}_+$ for each repositioning location $r \in R$ and each time $t \in T$. It specifies an expected cancellation value based on the estimated spatially and temporally proximate cancellations. The adaptive learning process applied to obtain reliable estimates of the expected cancellations c forms the core of our approach and is discussed in detail in the next section. To rank the different repositioning locations, based on this matrix c a min-max-normalized expected cancellation vector $\bar{c}_k$ for a decision epoch k can be extracted as shown in Equation (2):

$$\bar{c}_k = \frac{c_{rk} - \min(c_{rk})}{\max(c_{rk}) - \min(c_{rk})} \forall r \in R. \tag{2}$$

As a result, vector $\bar{c}_k$ specifies, on a scale of 0 to 1, the expected cancellations at time t_k for each repositioning location $r \in R$. Similar normalizations are performed for the two other base heatmaps to ensure comparable scaling when calculating the final heatmap.

While the first heatmap is based on learned information, the heatmaps of Figure 2(b) and (c) are based on state information S_k. Figure 2(b) is determined by the min-max normalized frequencies

with which each repositioning location $r \in R$ occurs in the vector of current and next repositioning locations of idle drivers ρ_k:

$$\bar{b}_k = \frac{|r \in \rho_k| - \min(|r \in \rho_k|)}{\max(|r \in \rho_k|) - \min(|r \in \rho_k|)} \forall r \in R. \tag{3}$$

Repositioning locations $r \in R$ most occupied by idle drivers are thus assigned a value of 1 and the least occupied ones a value of 0. Figure 2(c), on the other side, results from the min-max normalized travel times $\mathcal{T}(n_k, r)$ from current location n_k to the repositioning locations $r \in R$:

$$\bar{\theta_k} = \frac{\mathcal{T}(n_k, r) - \min(\mathcal{T}(n_k, r))}{\max(\mathcal{T}(n_k, r)) - \min(\mathcal{T}(n_k, r))} \forall r \in R. \tag{4}$$

The final repositioning heat map, shown in Figure 2(d), is derived according to Equation 5, where the signs emphasize that high values for $\bar{c}_k$ indicate high opportunities and for $\bar{b}_k$ and $\bar{\theta_k}$ low ones:

$$x_k = \bar{c}_k - \bar{b}_k - \bar{\theta}_k \tag{5}$$

Values $\bar{b}_k$ and $\bar{\theta}_k$ are derived directly from state S_k. Thus, the resulting policy directly depend on the parametrization of c. In the following, we will describe how c is approximated.

4.3. Adaptive Learning Process

In the presented heatmap design, the expected cancellations c play a crucial role since they reflect the complex interplay of future demand, its routing and assignment as well as future driver repositioning. The challenge in approximating them is that using the expected cancellations in the heatmap design results in shifts in driver distribution that may affect subsequent states and eventually lead to new cancellations in different areas of the city. We therefore propose a learning process that uses and adapts the approximated cancellations such that, as we show in our experiments, cancellations are stepwise reduced until they converge to a final matrix c.

In the following, we first give an overview of the learning process based on Figure 3 before explaining the algorithmic details. Figure 3 shows the data involved as well the four main steps of the learning process. The data are:

A: As input the expected demand $\mathbb{D}$. We assume that in practice it can be derived from forecasts or historical booking data. However, for our experiments, we use a pool of trip requests to sample expected demand scenarios.

B: The estimated expected cancellation values learned from the previous iteration. The values are updated after each iteration and used as the expected cancellation matrix c introduced before in the design of the policy applied in the next iteration.

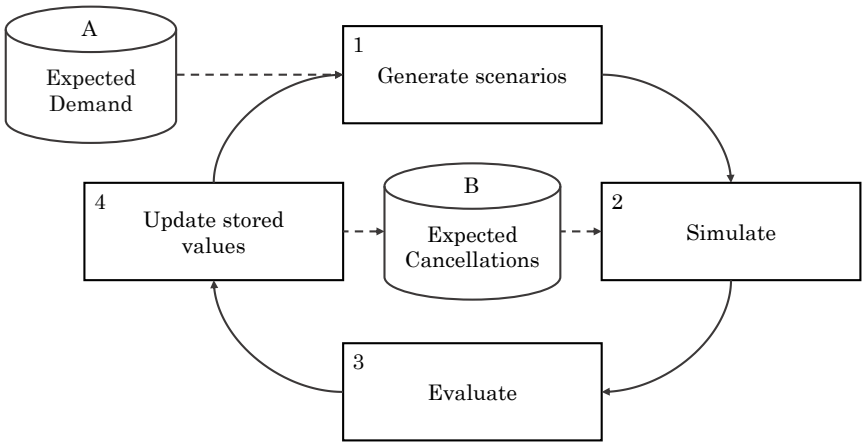


Figure 3 Adaptive learning process

Using this data, the following four process steps are executed iteratively:

1: Generation of multiple scenarios based on the expected demand $\mathbb{D}$. Each demand scenario depicts the demand for the entire service horizon. For this purpose, the demand is sampled (from A) on the basis of varying random seeds. This implies that the spatial and temporal distribution of demand remains fairly constant, but the actual trip requests vary.

2: Simulation of the generated scenarios. In the simulations, a repositioning heatmap is determined in each occurring decision epoch k based on the state information S_k and the expected cancellations c (stored in B).

3: The occurrence of cancellations in time and space is evaluated for all scenarios.

4: The obtained cancellations are then used to replace the stored expected cancellations (in B).

In the following, we will discuss the details of the algorithm using pseudocode 1. First, we motivate two important inputs: the temporal discount function $F(t)$ and the spatial discount matrix $\bar{g}$, as they control the extent to which a cancellation influences expectations in time and space:

- **Time:** With respect to the *temporal discount function* $F(t)$, preliminary experiments indicated that cancellations that occurred within the next 90 minutes should be accounted to varying degrees. The resulting function $F(t)$ increases for the first 20 minutes to reflect that repositioning is unlikely to prevent cancellations in the short run, and then decreases monotonically as a repositioned driver is more likely to be otherwise engaged in the long run. Cancellations later than $t + 90$ are not integrated in the cancellation calculation for time t.

- **Space:** Cancelled demand could have been satisfied not only from the closest but from several relocation locations. To integrate this, we define a *spatial discount matrix* $\bar{g}$ defined in Equation (6):

$$\bar{g} = (1 - \frac{(\mathcal{T}(r,l) - \min \mathcal{T}(r,l)}{\max(\mathcal{T}(r,l)) - \min(\mathcal{T}(r,l))} \forall l \in R) \forall r \in R. \tag{6}$$

This matrix provides the converted min-max normalized travel times between all repositioning locations $r \in R$. Thus, neglecting the temporal aspect, cancellations are factored with a value of 1 for the nearest repositioning location and with a value of 0 for the most distant one. The continuously decreasing spatial impact of cancellations favors the balancing of expected cancellations. Moreover, the indirect consideration of travel times also increases the value of conveniently repositioning locations to a greater extent.

In addition to the two discount functions, the input of the algorithm consists of the expected demand $\mathbb{D}$ for sampling scenarios, and the set of repositioning locations R. The output is a policy π_c that defines a repositioning opportunity heatmap for each state S_k based on the expected cancellation matrix c.

The algorithm is initialized in the first 4 lines. As line 3 indicates no cancellations are expected initially. Repositioning is thus restricted to the nearest repositioning location in the initial iteration.

In line 6 the actual learning process starts running over i^{max} iterations. Each iteration starts by deriving a heatmap generating policy π_c based on the current expected cancellations c (line 7). In addition, an empty set P is defined to collect cancellations to be processed later (line 8).

Line 11 to 16 cover the steps scenario generation, simulation and evaluation. How often these steps are executed depends on the number of scenarios j^{max} considered for each learning iteration. In line 12, a scenario D_j is generated from the expected demand $\mathbb{D}$ by sampling of trip requests (function $scenario(\mathbb{D})$). The simulation of the scenario D_j is performed next, applying the incumbent policy π_c (function $simulate(D_j, \pi_c)$). In step 14, the cancellations are transferred from the simulation result O_{D_j,π_c} to the set P (function $cancellations(O_{D_j,\pi_c})$).

The cancellations collected in P are subsequently used in lines 18 to 25 to update the expected cancellations c. For this purpose, it is iterated for each repositioning location $r \in R$ over each cancellation $p \in P$. Furthermore, for each combination of r and p it is iterated over each time $t \in T$ for which the function call $F(t - t^p)$ returns a positive value (line 21). The update of a matrix entry $c_r t$ is executed in line 22. In the update, an increase of the previous value of $c_r t$ is done by adding the result of multiplying the temporal discount factor $F(t - t^p)$ by the value of the spatial discount matrix $\bar{g}_{rl^p}$.

After the expected cancellations c have been updated, the next iteration of the learning process begins. For our experiments we set the number of iterations performed $i^{max} = 20$, as well as a number of simulated scenarios per iteration $j^{max} = 20$, as preliminary tests indicated that this amount is more then sufficient. We denote the final policy *Repositioning Opportunity Heatmaps* (ROH).

input : Expected demand $\mathbb{D}$, Repositioning Locations R
Temporal-Discount-Function $F(t)$,
Spatial-Discount-Matrix $\bar{g}$
output: Policy π_c

Function $ALP(\mathbb{D}, R, F, G)$
  $c(R, t^{\max}) \leftarrow 0$; `// Initialize cancellation matrix`
  $i \leftarrow 0$; `// Initialize learning iterations`
  `/* Perform` $i^{\max}$ `learning iterations: */`
  while $(i < i^{\max})$ **do**
    $\pi_c \leftarrow c$; `// Create policy from cancellation matrix`
    $P \leftarrow \emptyset$; `// Initialize empty set of cancellations`
    $j \leftarrow 0$; `// Initialize simulation iterations`
    `/* Simulate` $j^{\max}$ `scenarios: */`
    while $(j < j^{\max})$ **do**
      $D_j \leftarrow scenario(\mathbb{D})$; `// Generation of a scenario`
      $O_{D_j,\pi_c} \leftarrow simulate(D_j, \pi_c)$; `// Simulation of the scenario`
      $P \leftarrow P \cup cancellations(O_{D_j,\pi_c})$; `// Collecting the cancellations from the results`
      $j \leftarrow j + 1$
    end
    `/* Update entries in cancellation matrix: */`
    for $(r \in R)$ **do**
      for $(p \in P)$ **do**
        $t \leftarrow t^p$
        while $F(t - t^p) > 0)$ **do**
          $c_{rt} \leftarrow c_{rt} + (F(t - t^p) \cdot \bar{g}_{rl^p})$; `// Update of an entry in the cancellation matrix`
          $t \leftarrow t + 1$
        end
      end
    end
    $i \leftarrow i + 1$
  end
  return π_c

Algorithm 1: Adaptive Learning Process

5. Experimental Setup

In this section, we present the setup for the computational evaluation of the proposed ROH. We present the implementation of the decentralized repositioning decision-making, and the design of the test instances. We also introduce the repositioning policies that serve as performance benchmarks.

5.1. Instances

In the following, we describe how we designed the service area from real-world data, and how we create and assign demand of the ride-sharing system at hand.

Service Area. For the computational evaluation, we investigate ride-sharing systems operating in the urban area of New York City. More precisely, we analyze the performance of two individual systems that operate separately in the boroughs of Manhattan and Brooklyn. They differ in size, shape, and demand distribution, which should make the comparison interesting. Manhattan is comparatively smaller, characterized by its island shape, and shows a more temporally and spatially unbalanced demand. In contrast, Brooklyn is much broader and relatively circular, with a demand concentration north of the center, although demand appears to be more evenly distributed. In our simulations, the areas are represented by 3000 unique locations sampled from taxi trip data from January 2014 (NYC Taxi and Limousine Commission n.d.). Travel times are based on OpenStreetMap free-flow travel times which are multiplied by two to obtain a simple approximation of rush-hour traffic congestion. For the operation of the ride-sharing system, for each service area, 100 of the 3000 possible locations are defined as taxi rank-like repositioning locations. To achieve an fairly even distribution of these locations, the selection is made by means of a k-medians clustering algorithm using latitude and longitude values. The crowdsourced ride-sharing fleet is assumed to consist of 200 drivers operating homogeneous vehicles with four passenger seats.

Demand Creation. The planning period under consideration covers an 8-hour afternoon shift from 14:00 to 22:00. To exclude warm-up and cool-down phases of the simulation, the first and last 30 minutes are not taken into account for the evaluation. At the beginning of the planning period, the drivers are randomly distributed among the repositioning locations ready to fulfill incoming transportation requests. A number of 6400 incoming transportation requests is assumed per simulation run. Precomputational experiments revealed that a total of 6400 requests would allow for a reasonable relationship between fleet size and transportation requests, so that cancellation rates range from a minimum of 2% to a maximum of 30%. Due to the warm-up and cool-down phases, the number of requests within the evaluated period is not fixed, but varies around 5830. Each transportation request includes the transportation of one user. For each simulation run, transportation requests are sampled from a pool of 100,000 trips per service area performed by Uber or Lyft in September 2019 (NYC Taxi and Limousine Commission n.d.).

We obtain a realistic spatial and temporal distribution following the real demand as follows. Since information on request times is not provided in the data set and information on pickup and drop-off locations is only available at zone level, we interpret the pickup times as request times and randomly select the pick-up and drop-off locations of the sampled trips for each simulation

run from the subset of locations belonging to the respective zones. Each performed simulation run thus differs in the initial locations of the drivers, the request times, and the pick-up and drop-off locations of the trips. For the requesting users, it is assumed that a maximum waiting time of 10 minutes and a maximum travel time of 1.5 times the direct trip's travel time is acceptable.

Demand Assignment and Routing. The centralized assignment or cancellation of incoming transportation requests is subject to time and capacity-related constraints and is performed by means of a min-travel-time insertion heuristic. Given the limited number of repositioning locations, the FIFO principle is applied to equally located vehicles in order to achieve an equitable assignment. Finally, it is assumed that a repositioning that has been initiated must be completed, i.e., drivers do not check the app while driving.

5.2. Modeling Driver Compliance

In case of full driver compliance, the drivers would always select the relocation location $r \in R$ with maximum value x_{rk}. However, since the drivers are crowdsourced, they may not be fully compliant. To implement decision-making with varying levels of driver compliance, we assume that in addition to the heatmaps provided by the ride-sharing platform, each driver has his or her own idea of a heatmap for repositioning, capturing individual behavior and experience. We focus on the two most prominent factors, current demand and repositioning time (Urata et al. 2021). We consider one type of information at a time to analyze how varying compliance with the platform-based heatmaps affects the performance of the service and the concerned drivers.

Non-compliance with demand as indicated by the platform-based heatmaps is linked to the assumption that drivers may prefer to follow currently observable trends rather than relying on the expectations of the central platform provider. To generate this additional decentral driver heatmap, we use the information on how many times a repositioning location was the nearest to a pickup requested in the past 10 minutes. The decentral driver heatmap for non-compliance with recommended repositioning time is derived from the travel times between the current driver location and the repositioning locations.

Non-compliant decision-making is then implemented by combining information from the platform-based heatmaps with the decentral driver heatmap. The combination of both heatmaps requires min-max normalization of both heatmaps' scores. The driver heatmap scores are weighted to control the extent of non-compliance with the provider-based heatmap. After the combination, the repositioning location with the highest score is selected as the repositioning target. The weighting is determined for each driver, by sampling a normal distribution with a mean value equal to 0. Thus, on average, all drivers are fully compliant, while the actual non-compliance varies considerably depending on the standard deviation of this distribution.

We note that the sampled weighting for a driver can be positive or negative. This means that the decision-making does not only vary in the degree of non-compliance, but also in how the non-compliance affects the individual decision-making. We interpret positive and negative weights as follows. In terms of current demand, a positive weighting means that drivers partially ignore information from the provider-based heatmap in order to move closer to the demand center. We call this type of behavior "demand chasers". In contrast, a negative weighting implies that the provider-based heatmap is ignored to avoid the current demand center. We call this "demand gap chasers". Similarly, when repositioning times are positively weighted, the resulting non-compliance with the provider-based heatmap leads to longer repositionings. We call this type of behavior "time investors". For negative weights, shorter repositionings are preferred, which we call "time savers".

In the experiments, we will examine the implication of decreasing compliance on cancellation rates. As an extreme case, we set the standard deviation to zero leading to fully compliant drivers. Then, the standard deviation for determining the weighting per driver is exponentially increased from 0.15 to 19.2 across the experiments to model increasing non-compliance.

Furthermore, we will analyze whether drivers may benefit from decreasing compliance in their decision-making. For this analysis, a standard deviation of 1.2 is used to determine the weighting per driver. This value implies a balanced ratio of about 50% of the decisions being made in compliance with the provided heatmap.

For all experiments, we will use ROH iteratively learned under the assumption of compliant drivers. Re-learning with adapted non-compliant decision making has been tested for a few cases, but brought only marginal improvements. This is probably due to the fact that ROH becomes less important as compliance decreases, while at the same time increasing variance in driver decisions makes it more difficult to learn expected cancellations.

5.3. Benchmarks

In our computational experiments, we will compare varying levels of compliance with fully compliant drivers who always follow the heatmap as proposed by the service provider, and three further benchmarks policies: "Nearest Repositioning" (NR), "Model Predictive Control" (MPC), and "Expected Demand Heatmaps" (EDH). These three policies were selected as they feature different degrees of freedom with respect to repositioning decisions of idle drivers. NR specifies that a driver always repositions to the nearest repositioning location as soon as it becomes idle, which minimizes repositioning efforts and freedom. For MPC, we implemented a mixed-integer programming model proposed by Pouls, Meyer, and Ahuja (2020). This model is periodically solved for centralized repositioning of all idle drivers in order to maximize the coverage of the predicted demand at the minimal number of repositionings and minimum travel times. Details of how we

adapted this approach can be found in Appendix A.1. Finally, EDH reflects the case of minimal decision support for drivers' decentralized repositioning by indicating only the expected demand over the next 30 minutes. Non-compliant decision-making is considered for EDH in the same way as previously explained for ROH.

6. Results

We will present the results in two parts. In the first, we aim at investigating the effectiveness of the proposed heatmap design under "perfect" operating conditions, evaluating the impact of its application on the different actors of a ride-sharing system. For this purpose, we will assume compliant drivers who follow the repositioning recommendations provided by ROH. This is equivalent to a centralized variant of the ride-sharing system under study w.r.t. repositioning decision making. The evaluation will be made mainly in comparison to the three benchmark policies presented in Section 5.3. In the second part, we will analyze repositioning under varying compliance levels. Furthermore, we examine how the decentralized decision factors differ for individual drivers. All reported results represent the average over 100 simulation runs, with instances generated based on different random seeds for the preliminary experiments, the adaptive learning process, and the evaluation.

6.1. Evaluation with Compliant Drivers

In the following, we will provide an overall analysis of the results for compliant driver behavior. Based on this, we will demonstrate the learning process of the heatmaps and show their impact on service availability and driver fairness.

6.1.1. Objective Value. For the setting of decision-making with compliant drivers, we first analyze the impact of heatmap-based repositioning on minimizing cancellations. The idea is to demonstrate the performance of the proposed heatmap compared to the benchmark policies and to investigate the impact of the information considered to create ROH. The cancellation rates are shown in Figure 4 for the service areas of Manhattan and Brooklyn. The crosses below indicate the corresponding repositioning approach. We compare the proposed ROH to the benchmark policies NR and MPC as well as to ROH variants in which one of the information is omitted.

For both service areas, the proposed ROH yield the lowest cancellation rates of 2.2% and 5.7%, respectively. NR is clearly performing worst, with about 12.3% of transportation requests being canceled in Manhattan and 12.7% in Brooklyn. The second best results are obtained with MPC, with a slightly higher cancellation rate in Manhattan of 2.7% and a significantly higher one of 7.0% in Brooklyn. For the other ROH variants, in the case of Manhattan, neglecting the expected cancellation information proves especially disadvantageous, while in Brooklyn, neglecting repositioning

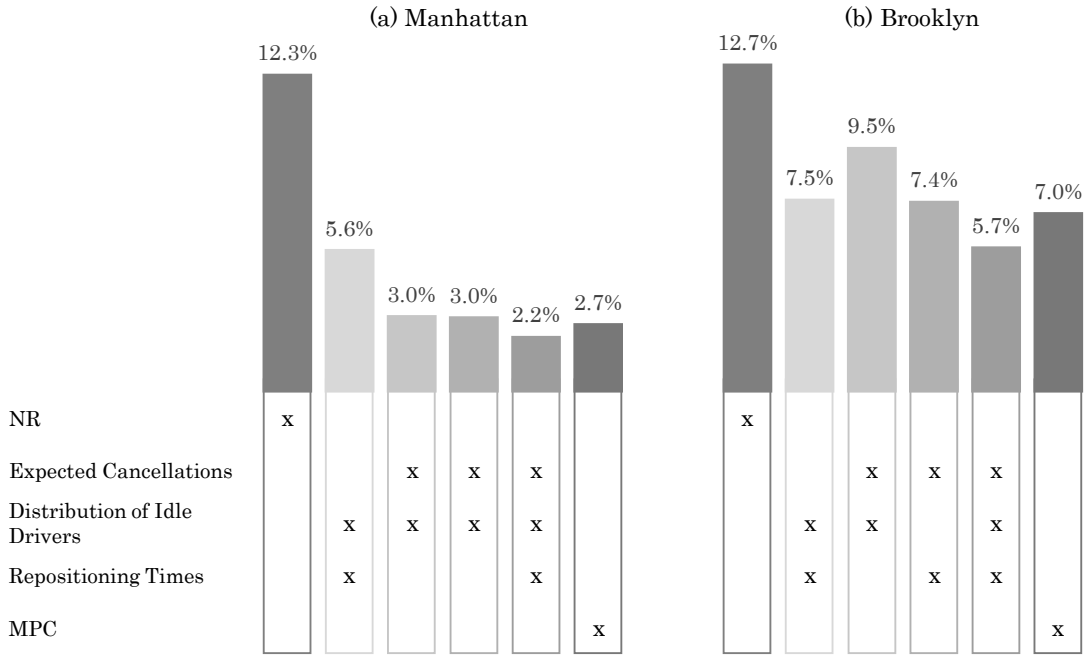


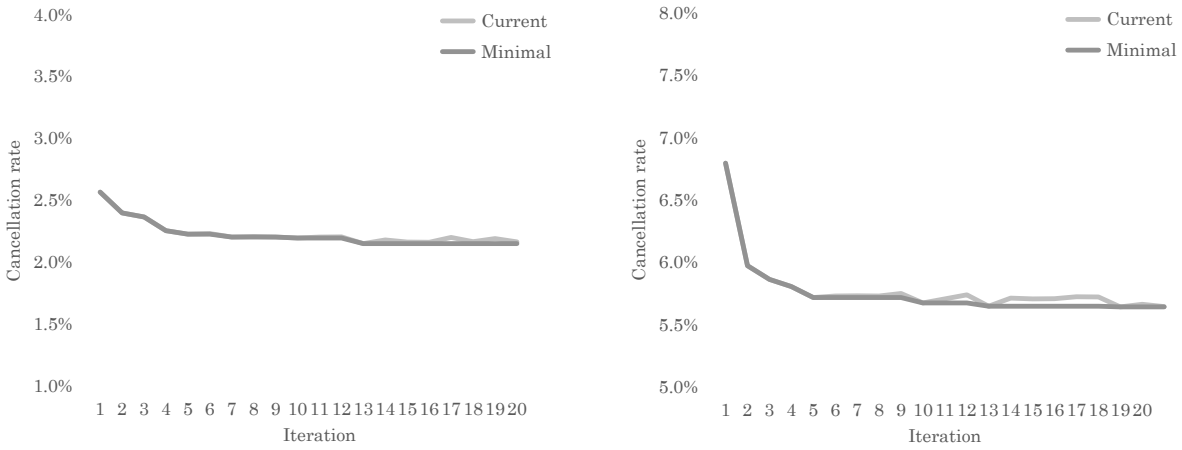
Figure 4 Cancellation rates

time leads to inferior results. Furthermore, the comparison with NR proves the high potential of active repositioning of unoccupied drivers. The resulting reduction in cancellations is significant in Manhattan and in Brooklyn. Moreover, it is plausible that in the rather small service area of Manhattan, dominated by a distinct demand center, *expected cancellations* is the most important information, whereas in the comparatively larger service area of Brooklyn, the importance of *repositioning times* dominates. Hence, the different characteristics of the service areas in terms of size, shape, and distribution of demand have a significant impact on the potential for repositioning idle drivers as well as on the challenge of its realization.

In summary, the performance of ROH is best when all three pieces of information are considered. This demonstrates that the proposed heatmaps are a powerful alternative for repositioning idle vehicles in a ride-sharing system with compliant drivers, even compared to a model predictive control approach.

6.1.2. Learning Process. One important feature of our policy is the adaptive learning of the cancellations. The corresponding learning curves for Manhattan and Brooklyn are shown in Figure 5, with the cancellation rate plotted on the X-axis and the process iteration plotted on the Y-axis. The light gray curve reflects the cancellation rate of the current iteration, the dark gray curve refers to the lowest cancellation rate found so far.

Iteration 1 corresponds to the cancellation rate after the initial expected cancellations are included in the ROH. For Manhattan, the progression of the curves, starting from an value of about



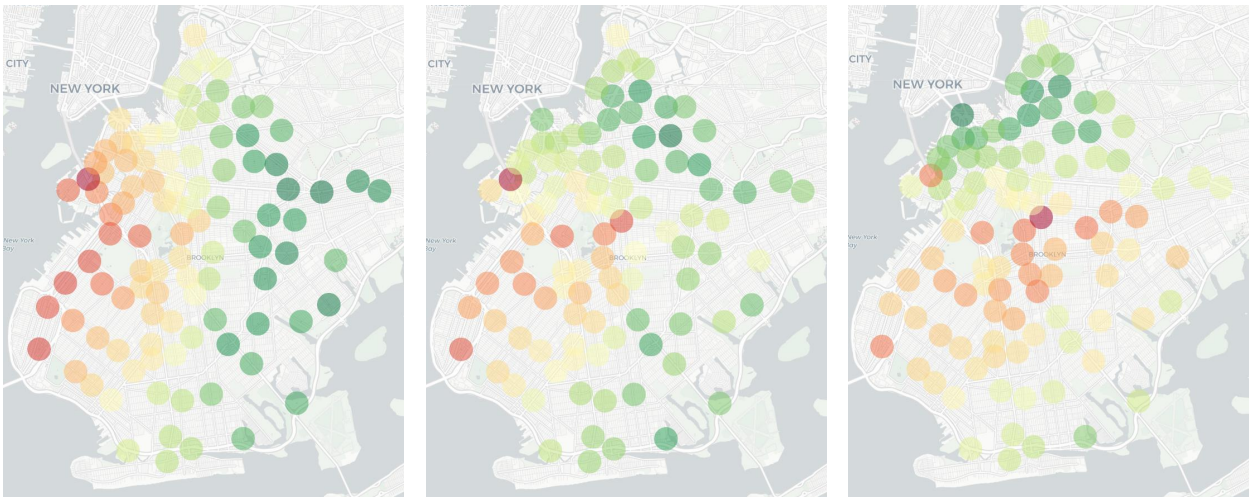
(a) Manhattan (b) Brooklyn

Figure 5 Learning Curves

2.6%, illustrates a steady decrease in cancellations to about 2.2% in the first five iterations. For all further iterations, the brighter curve indicates that the cancellation rates oscillate around 2.2%. For Brooklyn, the pattern looks quite similar, however with a much more distinct improvement in the first iterations. Here, the initial cancellation rate is about 6.8%, decreases to 5.7%, and then oscillates on this level.

The presented learning curves demonstrate that the iterative learning process improves the overall performance of the heatmaps. The substantial improvements are obtained after only a few iterations, so the expected cancellations converge very quickly. The noticeable difference between the levels of improvement in the areas suggests that in Manhattan, due to the focus on the middle of the island, the initial values are already quite dependable. In contrast, for Brooklyn, the more balanced demand structure appears to make it more critical to adjust expected cancellations through the learning process. **In conclusion, adaptive learning of expected cancellations can yield a significant improvement with relatively modest effort, especially in the case of more complex demand structures.**

In addition to presenting the general benefits of adaptive learning, we also want to illustrate the actual adaptations of the expected cancellations through exemplary heatmaps. For this purpose, we compare for a fixed point in time the distribution of expected cancellations after one and four iterations as well as after completion of the learning process. Figure 6 shows the corresponding heatmaps for 14:30, with the repositioning locations indicated as circles colored following the normalized expected cancellations from green (high) over yellow (medium) to red (low). The first heatmap (6a) shows a clear distribution characterized by decreasing expectations from the east (green) to the west (red) of the service area. After four iterations of the adaptive learning process, this distribution appears to be diminishing, with expected cancellations increasing particularly for



(a) 1. Iteration (b) 4. Iteration (c) Final

Figure 6 Exemplary Learning Process

repositioning locations in the north and decreasing for those in the east (6b). For the final heatmap, this trend is concluded with opportunities expected primarily in the north of the service area (6c).

From these exemplary visualizations, it can be reasoned that the distribution of expected cancellations systematically adapts over the course of the learning process, leading to better repositioning decisions that reduce cancellations.

6.1.3. Service Availability. In the evaluation of the learning process, it became visible that the distribution of expected cancellations is learned systematically. This poses the question of whether the associated avoidance of cancellations leads to an improved and more balanced service availability throughout the service area. This would be critical for user retention, as insufficient service availability or systematic discrimination against certain parts of the service areas lead to user dissatisfaction and churn.

To examine regional service availability, we analyze the cancellation rates per repositioning location compared to those of NR. For this purpose, the requests are assigned to the repositioning location that is nearest to the pickup location in terms of travel time. Figure 7 shows the corresponding cancellation rates per repositioning location by colored circles as well as their demand volumes by means of their size. Here, the scale on which the cancellation rate is assigned to color ranges from a minimum of 0% (dark blue) to a maximum of 50% (dark red). Therefore, a small blue circle, for instance, indicates that a low cancellation rate for a low demand volume is observed in the vicinity of the repositioning location, while a large red circle would indicate a high cancellation rate for a high demand volume.

We first focus on the results for Manhattan, shown in the top two figures. With regard to the distribution of demand, i.e. the size of the circles, the demand center in the middle of the island is

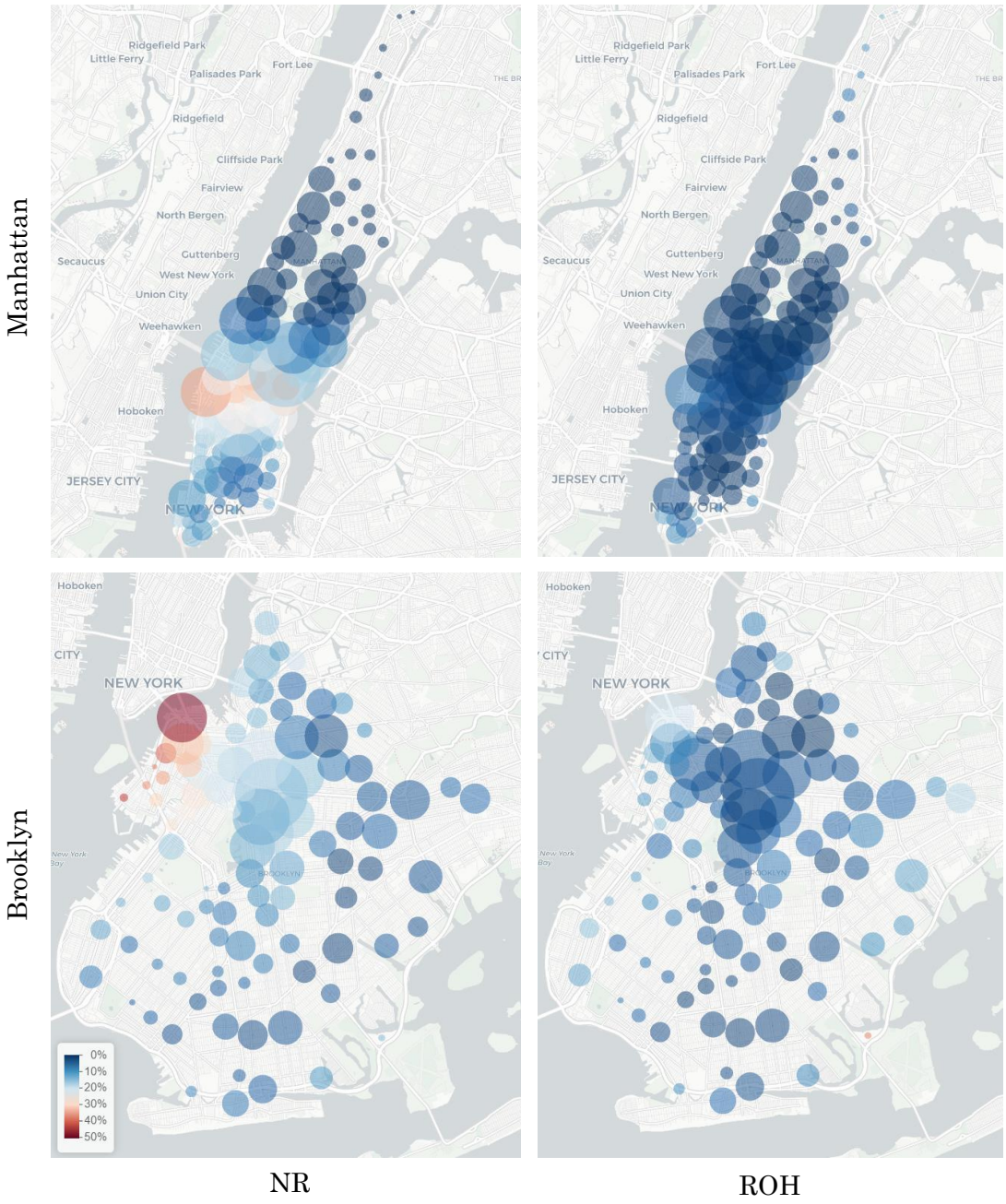


Figure 7 **Distribution of cancellation rates (color) and demand (size)**

clearly visible as well as a relatively large region with very low demand in the north. Regarding the color-visualized cancellations, for NR, the light red circles in the lower part of the demand center are noticeable, indicating increased cancellation rates. Furthermore, lower cancellation rates are visualized in the northern region, even in the far north where demand is very scarce. For ROH, all repositioning locations are colored in different intensities of blue, indicating relatively balanced cancellation rates. However, the upper and lower corners and parts of the demand center appear to be slightly brighter blue, indicating a moderate increase in the cancellation rates in these regions.

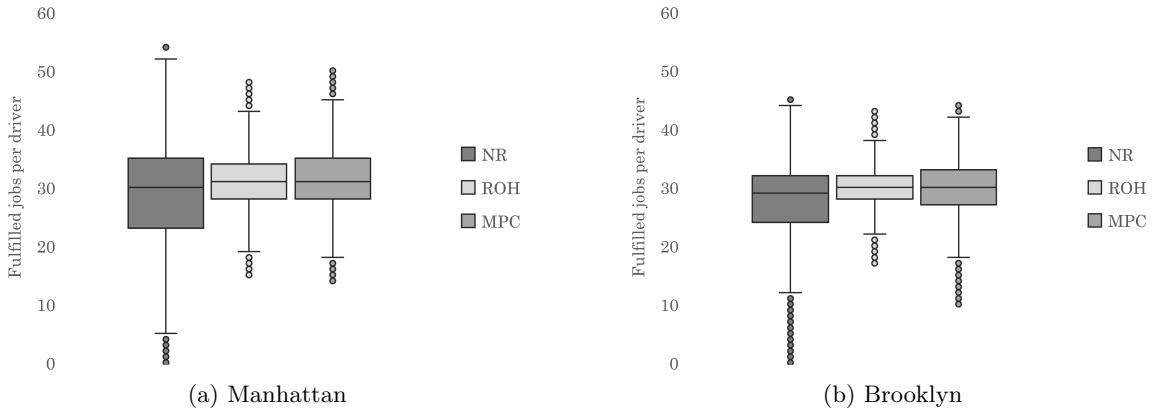
For Brooklyn, the demand center appears slightly north of the geographic center. From the demand center, the volume of demand decreases slightly towards the borders, with isolated repositioning locations with hardly any demand in the vicinity. Cancellation rates follow a pattern similar to the one observed in Manhattan. For NR, the light to dark red circles north of the demand center are particularly prominent, indicating comparatively high cancellation rates in this region. Again, in the case of ROH, all circles are in the blue range of the scale and thus indicate rather low cancellation rates, with one small exception in the southeast where the size of the circle indicates a very scarce demand.

From the results of both service areas, it can be concluded that ROH greatly contribute to increased and more balanced availability of service compared to NR. It is apparent that the heatmaps particularly help to decrease the otherwise high cancellation rates in high-demand regions. Moreover, these improvements are only slightly detrimental to regions with very scarce demand located at the outermost corners of the service area. The balancing improvement can be attributed to better learning of expected earning opportunities while focusing on the repositioning of idle vehicles ensures that only underutilized resources are redistributed. **In conclusion, ROH offer providers the opportunity to serve more users with a constant pool of drivers and to ensure a more balanced service availability across their service area.** Moreover, these enhancements are likely to increase user satisfaction and therefore their commitment to the ride-sharing system.

6.1.4. Driver Fairness. After showing that platform providers would benefit from the application of ROH, it remains to be investigated whether the same is true for drivers.

To evaluate driver satisfaction, we focus on the distribution of jobs per driver, which is a proxy for the distribution of earnings. The corresponding boxplots for Manhattan and Brooklyn are shown in Figure 8 for ROH in comparison to those of NR and MPC. Each of these boxplots represents for $200 \text{ daily drivers} \times 100 \text{ simulations} = 20,000$ drivers the number of fulfilled transportation requests.

For Manhattan, the median of transportation requests is about 30 for all approaches, with slightly higher values for ROH and MPC as they lead to fewer cancellations. More distinct differences can be observed among the interquartile ranges. This range is clearly widest for NR with 23 to 35 jobs, decreases to 28 to 35 jobs for MPC, and again slightly to 23 to 34 jobs for ROH. The same ranking can be observed with respect to the length of the whiskers. For NR, outliers reveal that some drivers perform only a very small number of jobs. In contrast, the outliers in ROH and MPC are less pronounced and occur both positively and negatively. For Brooklyn, overall, the results are very similar to Manhattan except that the interquartile ranges and the whiskers are significantly narrower for all three approaches. In return, however, more and very pronounced negative outliers are visible for NR and MPC.



(a) Manhattan (b) Brooklyn

Figure 8 Number of fulfilled jobs per driver

The decreasing magnitudes of interquartile ranges illustrate that the distribution of jobs among drivers is more balanced for ROH compared to NR and, to a lesser extent, to MPC. This involves far fewer drivers receiving a significantly below-average and above-average number of jobs. The absence of extremely negative outliers in ROH indicates that, in contrast to the other approaches, no driver has to suffer from major drawbacks when following the repositioning recommendations of our heatmaps. Another interesting aspect is that independent of the approach under consideration, Brooklyn tends to provide a more balanced distribution of jobs per driver than Manhattan. Reasons may include the larger size of Brooklyn as well as the more even distribution of demand; this combination probably ensures that drivers are spread out further across the serves area and still get assignments. **In summary, ROH contribute to a more balanced and thus fairer distribution of jobs among drivers.** However, a more even distribution of earnings can have both positive and negative effects concerning a driver's personal income. Thus, the extent to which the equitable distribution contributes to the acceptance of the proposed heatmaps depends on prior earnings as well as the driver's personal mindset toward equal opportunities. Regardless of the heatmaps, it appears that a balanced distribution of jobs per driver tends to occur more naturally in Brooklyn, while it must be maintained by the ride-sharing platform in Manhattan. Therefore, it may be of interest for service areas like Manhattan to address income balance directly when making repositioning recommendations.

6.2. Evaluation with Non-compliant Drivers

In the following, we examine what happens when drivers do not (completely) follow recommendations indicated by their heatmaps.

6.2.1. Provider Perspective. We first focus on the perspective of the provider of a ride-sharing platform and evaluate the impact of decreasing compliance of drivers on minimizing cancellations. Of the two factors presented in Section 5.2, in Figure 9, we analyze the effects of decreasing

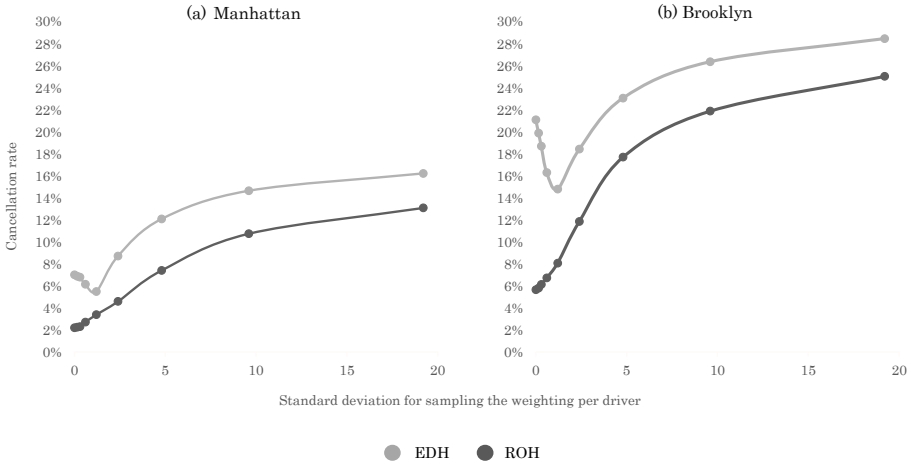


Figure 9 Demand: Cancellation rates at decreasing driver compliance

compliance with displayed information in favor of current demand information. The X-axis represents decreasing compliance by increasing standard deviations of weighting current demand in drivers' decentral decision-making, and the Y-axis indicates the system-wide cancellation rates for Manhattan (a) and Brooklyn (b), respectively. The curves show the results for EDH in light gray and for ROH in dark gray.

For Manhattan, ROH lead to smaller cancellation rates than EDH. For the latter, the cancellation rate initially decreases from around 7.0% to 5.5% before it increases up to 16.2% with decreasing compliance. For ROH, the cancellations also increase consistently with decreasing compliance from the initial 2.2% to about 13.1%. For Brooklyn, similar trends can be observed, although characteristics are more distinct there. For instance, the decrease in cancellations at EDH is initially stronger, yet the overall gap in cancellations compared to ROH is stronger as well. Furthermore, with decreasing compliance, cancellations increase more rapidly. One potential reason for this development is that demand is more evenly distributed in Brooklyn. This requires a more balanced distribution of drivers in the relatively large area compared to Manhattan where demand is clustered in a smaller area. As non-compliance reduces the balanced distribution, cancellation rates in Brooklyn rise more rapidly.

When drivers tend to ignore provided information in favor of current demand information, holistic repositioning heatmaps still perform better than expected demand heatmaps with respect to minimizing cancellation rates. However, EDH even benefit sometimes from non-compliance. This is because EDH only map one piece of information, while ROH take a more holistic view of the current decision state into account. Thus, for EDH, for smaller variances, using additional information can help to improve the otherwise one-dimensional distribution of repositioning drivers. In contrast, for ROH, a balanced distribution is already reflected in the design of the heatmaps.

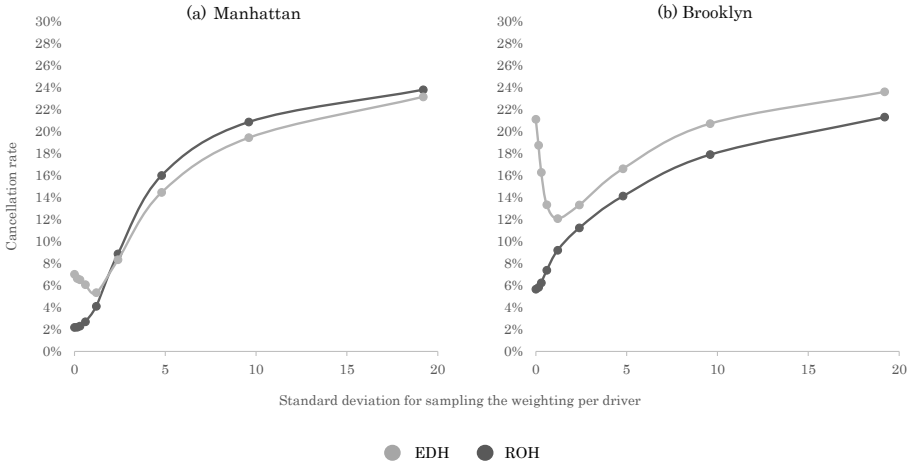


Figure 10 Repositioning time: Cancellation rates at decreasing driver compliance

Favoring current demand information has different implications for different service areas. As can be seen from Figure 9, especially for small levels of non-compliance, performance gaps are quite different. This is less detrimental for Manhattan and more for Brooklyn, which is consistent with the observations made in the previous section that making good repositioning decisions depends more on the expected cancellations in Manhattan, i.e. on the expected distribution of demand, and more on the repositioning times in Brooklyn.

Secondly, we examine the impact of ignoring heatmap information in order to save/invest repositioning time. Figure 10 shows the cancellation rates (Y-axis) in dependency of the applied standard deviation of the repositioning time weight (X-axis), which represents decreasing compliance with the heatmaps. Again, the results are illustrated with EDH in light gray and for ROH in dark gray for Manhattan (a) and Brooklyn (b).

For Manhattan, with decreasing compliance in order to save/invest repositioning time, cancellations at EDH first decrease from 7.0% to 5.4%, followed by an increase to 23.1%. For ROH, cancellations increase moderately at first, followed by a more rapid increase than for EDH. For Brooklyn, the characteristics of the results again differ slightly. Here, for EDH, a moderate non-compliance initially leads to a substantial decrease in cancellations. However, the subsequent increase in cancellations follows the continuous increase observed with ROH, yet on a higher level.

The results indicate that under most conditions, ROH lead to fewer cancellations than EDH. For Manhattan, ROH are more affected by non-compliant drivers, which is because ROH already covers repositioning time. Furthermore, EDH even benefit from moderately non-compliant drivers, while the cancellation rates for ROH increase consistently. Here, it seems that adaptive repositioning recommendations are already so well-designed that any deviation is detrimental. Decreasing

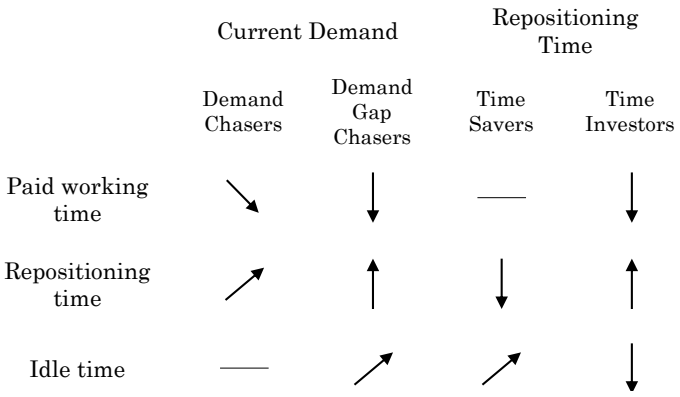


Figure 11 ROH: Correlations between time usage and decreasing driver compliance

compliance with heatmaps has a greater impact in Manhattan. Moreover, the performance of ROH is best when drivers always make decisions in compliance with the repositioning recommendations.

The extent to which non-compliance is detrimental depends on the combination of alternatively used information and service area. In Manhattan, drivers not complying in favor of current demand reduce system performance slightly less strongly. This is similar for Brooklyn, where those who do not comply with respect to repositioning times reduce system performance slightly less strongly. However, the overall rapid increase in cancellation rates indicates that it may be worthwhile to partly hire drivers to ensure a limited influence of decentralized decisions, even if this entails higher personnel costs (Lee and Savelsbergh 2015).

6.2.2. Driver Perspective. Lastly, we evaluate the impact of non-compliant repositioning decisions from a driver's perspective, considering the four types of individual driver behavior represented by demand chasers, demand gap chasers, time investors, and time savers, as described in Section 5.2. To this end, a working day is assumed to be covered by paid working time, in which assigned trips are performed, as well as idle and repositioning time. We compute Spearman's correlation coefficients between the amount of time spent per activity and the fixed weighting of the decentral driver heatmap. The resulting correlations for both service areas are illustrated by the symbols in Figure 11. The symbols can be interpreted as follows: "up" indicates a positive correlation above 0.5, "diagonally up" indicates a slightly positive correlation above 0.25, a dash indicates a low correlation between 0.25 and -0.25, "diagonally down"' indicates a slightly negative correlation below -0.25, and "down" a negative correlation below -0.5.

For demand chasers, the first column illustrates a slightly negative correlation between weighting and paid working time. In contrast, a slightly positive correlation is indicated for the repositioning time, while no correlation is observed for the amount of idle time. **This indicates that demand chasers make less earnings and spend more time repositioning than drivers who follow**

the platform-based heatmap. This can be attributed to the fact that regardless of their location, the corresponding drivers spend much time reaching the demand center, while eventually competing for assignments there with similarly acting drivers.

The second column demonstrates consequences for demand gap chasers. We can see a strong negative correlation between weighting and paid work time, as well as a strong and weak positive correlation with repositioning and idle time, respectively. Therefore, **for demand gap chasers, with increasing avoidance of the current demand center, drivers' earnings decrease and the time spent repositioning or idling increases.** This suggests that demand gap chasers are likely repositioning to regions with relatively low demand, only to perform another repositioning after a period of idleness.

The third and fourth columns illustrate the impact of non-compliance with information on repositioning times (time investors, time savers). The arrows for repositioning time indicate strong opposite correlations. In accordance with the positive and negative correlations of the repositioning time, there is a slightly positive and strongly negative correlation for idle time. Of particular interest, therefore, is the correlations for the paid working time. Here, the line indicates no significant correlation between the degree of repositioning avoidance and paid work time. **This means that time-savers make, on average, the same earnings as drivers who follow the recommendations of the platform-based heatmap.** The previous results, which showed a substantial increase in cancellations for NR demonstrate that time saving is only a good option if enough other drivers are repositioning. **In the opposite case of time investors, we see a strong negative correlation between weighting and paid work time, indicating that increasing repositioning beyond the recommendations leads to decreasing earnings.**

In conclusion, our results demonstrate that decreasing compliance with heatmaps mostly results in decreasing earnings and increasing repositioning times. An exception to this is the behavior of time savers, which can be beneficial for drivers in order to decrease repositioning effort and to increase the more convenient idle time. However, in order to obtain these benefits without decreasing earnings, a sufficient number of drivers must comply with the repositioning information recommended by the platform-based heatmaps. In this context, both game-theoretic considerations and platform-based repositioning incentives are of interest for future research.

7. Final Remarks

As we have illustrated how heatmaps can lead to improved operations for service providers, drivers, and customers, there are several avenues for future research. First, our experiments have shown that carefully designed heatmaps reduce service cancellations even in cases that drivers are less compliant with the provider's recommendation. Future research may focus on explicitly identifying

such non-compliant behavior by analyzing the drivers' previous decisions. This analysis could then be used to adapt the heatmap design accordingly, e.g., by providing "time savers" reluctant to leave their neighborhood with recommendations nearby and use "time investors" to cover areas further away. However, as we have seen in our experiments, heterogeneous driver behavior already leads to imbalances in their earnings even when they are treated equally by the service provider. The imbalances may increase in case the provider further differentiates driver preferences. This leads to the question of fairness as very picky drivers may get very lucrative jobs while others do the heavy lifting. Furthermore, such differentiation may increase the number of drivers gaming the system once they realize that their behavior influences their recommendations. Future research may therefore focus on a fair and balanced repositioning given the heterogeneous driver preferences and ways to disengage drivers from gaming the system.

Another interesting aspect of our experiments is that the impact of driver compliance differs for services areas with different spatio-temporal characteristics. While in the rather small area of Manhattan, even less compliant drivers lead to a good demand coverage, in the larger area of Brooklyn, non-compliance results in a rapid increase in cancellations. Thus, for larger and more "complex" service areas such as Brooklyn, future research may focus on a better balance when drivers are not fully compliant. Potential options could be different driver compensation in different areas of the city. Alternatively, the provider might complement the crowdsourced fleet with dedicated drivers. Finding the right balance in the fleet and using the potentially more expensive dedicated drivers effectively may be an interesting challenge for future research.

Further, even though our method is independent of the underlying assignment and routing strategy as well as of the drivers job acceptance and working time behavior, future work may analyze the value of jointly learning heatmap design and assignment strategies, potentially under driver non compliance with respect to job assignments. One (of many) challenges will be maintaining the explainability of the heatmap design and creating intuitive and effective assignment policies. Finally, in our experiments, we have shown that heatmaps lead to more and more fairly distributed earnings amongst the drivers and better service availability for the customers. Future research may analyze the longer-term impact of these improvements in comparison with the status quo., e.g., with respect to customer retention, business growth, and the drivers' trust in and adoption of the heatmap system.

Acknowledgments

Marlin Ulmer's work is funded by the DFG Emmy Noether Programme, project 444657906. We gratefully acknowledge their support.

References

Alnaggar A, 2021 *Optimization under Uncertainty for E-retail Distribution: From Suppliers to the Last Mile*. Ph.D. thesis, URL http://hdl.handle.net/10012/17336.

Ausseil R, Pazour JA, Ulmer MW, 2022 *Supplier menus for dynamic matching in peer-to-peer transportation platforms. Transportation Science* .

Bimpikis K, Candogan O, Saban D, 2019 *Spatial pricing in ride-sharing networks. Operations Research* 67(3):744–769.

Braverman A, Dai JG, Liu X, Ying L, 2019 *Empty-car routing in ridesharing systems. Operations Research* 67(5):1437–1452.

Conger K, 2021 *Prepare to pay more for Uber and Lyft rides*. URL https://www.nytimes.com/article/uber-lyft-surge.html.

Dayarian I, Savelsbergh M, 2020 *Crowdshipping and same-day delivery: Employing in-store customers to deliver online orders. Production and Operations Management* 29:2153–2174.

Dholakia UM, 2015 *Everyone hates Uber's surge pricing – here's how to fix it*. URL https://hbr.org/2015/12/everyone-hates-ubers-surge-pricing-heres-how-to-fix-it.

Ermagun A, Stathopoulos A, 2018 *To bid or not to bid: An empirical study of the supply determinants of crowd-shipping. Transportation Research Part A: Policy and Practice* 116:468–483.

Gdowska K, Viana A, Pedroso JP, 2018 *Stochastic last-mile delivery with crowdshipping. Transportation Research Procedia* 30:90–100.

Goncharova M, 2017 *Ride-hailing drivers are slaves to the surge*. URL https://www.nytimes.com/2017/01/12/nyregion/uber-lyft-juno-ride-hailing.html.

Guda H, Subramanian U, 2019 *Your Uber is arriving: Managing on-demand workers through surge pricing, forecast communication, and worker incentives. Management Science* .

Holler J, Vuorio R, Qin Z, Tang X, Jiao Y, Jin T, Singh S, Wang C, Ye J, 2019 *Deep reinforcement learning for multi-driver vehicle dispatching and repositioning problem*. Wang J, Shim K, Wu X, eds., *19th IEEE International Conference on Data Mining*, 1090–1095 (Piscataway, NJ: IEEE).

Hu B, Hu M, Zhu H, 2021 *Surge pricing and two-sided temporal responses in ride hailing. Manufacturing & Service Operations Management* .

Iglesias R, Rossi F, Wang K, Hallac D, Leskovec J, M Pavone, 2018 *Data-driven model predictive control of autonomous mobility-on-demand systems. 2018 IEEE International Conference on Robotics and Automation (ICRA)*, 6019–6025.

Jiao Y, Tang X, Qin Z, Li S, Zhang F, Zhu H, Ye J, 2021 *Real-world ride-hailing vehicle repositioning using deep reinforcement learning. Transportation Research Part C: Emerging Technologies* 130:103289, URL https://www.sciencedirect.com/science/article/pii/S0968090X21003004.

Lee A, Savelsbergh M, 2015 *Dynamic ridesharing: Is there a role for dedicated drivers? Transportation Research Part B: Methodological* 81:483–497.

Lei Z, Qian X, Ukkusuri SV, 2020 *Efficient proactive vehicle relocation for on-demand mobility service with recurrent neural networks. Transportation Research Part C: Emerging Technologies* 117:102678, URL https://www.sciencedirect.com/science/article/pii/S0968090X20305933.

Li B, Zhang D, Sun L, Chen C, Li, Shijian, Qi, Guande, Yang Q, 2011 *Hunting or waiting? discovering passenger-finding strategies from a large-scale real-world taxi dataset. 2011 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, 63–68.

Li X, Gao J, Wang C, Huang X, Nie Y, 2021 *Driver guidance and rebalancing in ride-hailing systems through mixture density networks and stochastic programming. 2021 IEEE International Smart Cities Conference (ISC2)*, 1–7.

Liu C, Chen CX, Chen C, 2021 *Meta: A city-wide taxi repositioning framework based on multi-agent reinforcement learning. IEEE Transactions on Intelligent Transportation Systems* 1–6.

Möhlmann M, Zalmanson L, Henfridsson O, Gregory R, 2020 *Algorithmic management of work on online labor platforms: When matching meets control. MIS Quarterly* .

NYC Taxi and Limousine Commission, n.d. *Trip record data.* URL https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page.

Pavone M, Smith SL, Frazzoli E, Rus D, 2012 *Robotic load balancing for mobility-on-demand systems. The International Journal of Robotics Research* 31(7):839–854.

Pouls M, Meyer A, Ahuja N, 2020 *Idle vehicle repositioning for dynamic ride-sharing.* Lalla-Ruiz E, Mes M, Voß S, eds., *Computational Logistics*, 507–521, Theoretical Computer Science and General Issues (Cham: Springer International Publishing and Imprint: Springer).

Powell JW, Huang Y, Bastani F, Ji M, 2011 *Towards reducing taxicab cruising time using spatio-temporal profitability maps.* Pfoser D, ed., *Advances in spatial and temporal databases*, 242–260, Lecture Notes in Computer Science (Heidelberg: Springer).

Rai HB, Verlinde S, Macharis C, 2021 *Who is interested in a crowdsourced last mile? a segmentation of attitudinal profiles. Travel Behaviour and Society* 22:22–31.

Savelsbergh M, Ulmer MW, 2022 *Challenges and opportunities in crowdsourced delivery planning and operations. 4OR* .

Sayarshad HR, Chow JY, 2017 *Non-myopic relocation of idle mobility-on-demand vehicles as a dynamic location-allocation-queueing problem. Transportation Research Part E: Logistics and Transportation Review* 106:60–77, URL https://www.sciencedirect.com/science/article/pii/S1366554517300121.

Spielman F, 2021 *Alderman accuses Uber, Lyft of 'predatory fares,' wants price cap imposed.* URL https://chicago.suntimes.com/city-hall/2021/5/24/22451667/uber-lyft-ride-share-hailing-surge-pricing-cap-city-council-ordinance-alderman-reilly-taxi-cabs.

Ulmer M, Savelsbergh M, 2020 *Workforce scheduling in the era of crowdsourced delivery. Transportation Science* 54:1113–1133.

Urata J, Xu Z, Ke J, Yin Y, Wu G, Yang H, Ye J, 2021 *Learning ride-sourcing drivers' customer-searching behavior: A dynamic discrete choice approach. Transportation Research Part C: Emerging Technologies* 130:103293.

Wallar A, van der Zee M, Alonso-Mora J, Rus D, 2018 *Vehicle rebalancing for mobility-on-demand systems with ride-sharing. 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 4539–4546 (IEEE).

Wang J, Shim K, Wu X, eds., 2019 *19th IEEE International Conference on Data Mining: 8-11 November 2019, Beijing, China : proceedings* (Piscataway, NJ: IEEE).

Wen J, Zhao J, Jaillet P, 2017 *Rebalancing shared mobility-on-demand systems: A reinforcement learning approach. 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, 220–225.

Xi J, Zhu F, Chen Y, Lv Y, Tan C, Wang F, 2021 *Ddrl: A decentralized deep reinforcement learning method for vehicle repositioning. 2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, 3984–3989.

Yu Z, Hu M, 2021 *Deep reinforcement learning with graph representation for vehicle repositioning. IEEE Transactions on Intelligent Transportation Systems* 1–14.

Yuan J, Zheng Y, Zhang L, Xie X, Sun G, 2011 *Where to find my next passenger*. Landay J, Shi Y, Patterson DJ, Rogers Y, Xie X, eds., *Proceedings of the 13th international conference on Ubiquitous computing - UbiComp '11*, 109 (New York, New York, USA: ACM Press).

Zhang R, Pavone M, 2016 *Control of robotic mobility-on-demand systems: A queueing-theoretical perspective. The International Journal of Robotics Research* 35(1-3):186–203.

Zhang R, Rossi F, Pavone M, 2016 *Model predictive control of autonomous mobility-on-demand systems. 2016 IEEE International Conference on Robotics and Automation (ICRA)*, 1382–1389.

Appendix

A.1. Details of MPC Benchmark

Even though the implementation of MPC is based on Pouls, Meyer, and Ahuja (2020), there are some differences. The first concerns the representation of the service area: To ensure comparability with the other policies, we define subareas for each repositioning location instead of using a grid. Thereby, each location belongs to the area of the nearest repositioning location in terms of travel time. Another difference is the prediction of the expected demand. Instead of reactive or perfect demand, we determine the expected demand by taking the average of the demand that occurred in preliminary simulations. Finally, there is a difference in the execution of the repositionings determined by the model. This has been adapted so that drivers always move directly to the repositioning locations, while the assignment of which driver performs which repositioning is performed by a second travel time minimizing model. Regarding the extensive parameterization of the first model, the original values are used, except for the two parameters that differ depending on the service area. These two parameters which indicate how many requests can be fulfilled between two repositioning periods and the distance over which a repositioning location can cover the demand of another repositioning location were tuned sequentially for Manhattan and Brooklyn as part of preliminary simulations.

Otto von Guericke University Magdeburg
Faculty of Economics and Management
P.O. Box 4120 | 39016 Magdeburg | Germany

Tel.: +49 (0) 3 91 / 67-1 85 84
Fax: +49 (0) 3 91 / 67-1 21 20

www.fww.ovgu.de/femm

ISSN 1615-4274