Paper ID #

Addressing Spatial Service Provision Equity for Pooled Ride-Hailing Services through Rebalancing

Tilmann Schlenther[1*], Gregor Leich[1], Michal Maciejewski[1], Kai Nagel[1]

*Corresponding author: schlenther@vsp.tu-berlin.de

1. Technische Universität Berlin, Chair of Transport Systems Planning and Transport Telematics, Straße des 17. Juni 135, 10623 Berlin, Germany

Abstract

Ride-hailing services have the potential to play a major role within future transport systems. While the operation as well as environmental, social and economic consequences are research subjects, the focus in the literature lies on system efficiency and profitability. In this study, vehicle relocation strategies are investigated with respect to service provision equity using an agent-based simulation approach. The aim is to identify ways of operation that support social fairness while maintaining the profitability and effectiveness of the service. The results show that a typical demand-anticipatory relocation strategy leads to an imbalance in the spatial distribution of mean wait times for the use case of Berlin. This can be significantly improved by applying different relocation strategies. As a consequence, the overall demand for the system rises while the ratio of empty vehicle kilometres increases slightly.

Keywords:

Ride Sharing and Hailing Services, Vehicle Repositioning

1. Introduction

Ride-Hailing services have been promoted to offer a modern type of mobility and potentially decrease the overall number of vehicles in the traffic system. This could have further beneficial effects such as possibly less air pollution [15, 24] or improved traffic flow if automated vehicles are operated [19].
Most studies focus on the optimization of operational parameters such as the fleet size or the user cost structure [16, 15, 5, 2] or on efficient algorithms to the Vehicle Routing Problem [27, 28]. Another, relatively smaller, focus lies on the integration with public transport [26, 30, 20, 10].
While profit-oriented operators would tend to limit their service area to districts of high demand, mostly located within city centers, a ride-hailing service that is considered to be a part of the public mobility offer would be most beneficial in areas with a coarse public transport network (i.e. suburbs or rural areas) [4, 15]. One challenge of offering such a service is to maintain reasonable waiting times in those sparsely populated areas, as the demand level can be relatively low compared to the city center. Waiting times are an important part of the perceived attractiveness of the transport mode and its overall travel time [18]. They are strongly influenced by the spatial distribution of idle vehicles, which in turn can be adjusted by vehicle rebalancing. Most rebalancing algorithms focus on optimizing demand throughput

or deal with systems where pooling is not considered [29, 3, 8, 1]. Winter et al. [29] investigated different rebalancing strategies in a non-pooled system under parking constraints for Amsterdam. The authors state that rebalancing might lower the system performance in terms of overall mean user wait time, empty mileage and service provision equity. However, the study is conducted with a fixed demand and does not take mode choice behavior into account. Whether rebalancing is beneficial might also depend on the target service quality level [13, 21].

Bischoff and Maciejewski [6] proposed a rebalancing strategy that relocates vehicles so that their spatial distribution mimics the spatial distribution of demand in the near future. This demand-anticipatory type of rebalancing strategy can lead to the effect that vehicles are rather put into areas of high demand or high population density. Areas of relatively low demand receive less idle vehicles, leading to higher waiting times which again lowers the demand. As a result, areas of low demand or low population density are served with poor service quality and the demand may die out, leading to an imbalanced system [29, 4, 22]. Spieser et al. [25] have further described the conflict of interest between customers, who "favor large fleets and frequent rebalancing", and the operator, who wants to maximize the profit. This paper investigates rebalancing strategies that are designed for a public-transport-like ride-hailing service. The goal is to maintain waiting time equality throughout the entire service area to support the social fairness of the system. In order to account for the operator's perspective, the resulting demand is supposed to be at least as high as in comparable scenarios using an existing rebalancing approach or no rebalancing, respectively.

2. Methodology

This section provides background information on the topic of rebalancing as an optimization problem as well as the formulation of several targets for the solution and corresponding mathematical functions which are based on an existing approach. In the following, we explain the setup of simulations and the performance indicators which are the basis for the conducted analysis.

2.1 Rebalancing as an optimization problem and rebalancing target

In this study, we model vehicle rebalancing as a transportation problem [11] where we want to find minimum cost vehicle flows from the zones with surplus of vehicles to the zones with deficit of vehicles. Because the state of the demand-responsive transport (DRT) system is continuously changing on both the demand and supply side, the optimum relocation of idle vehicles needs to be regularly recalculated. Therefore, we solve the transportation problem at a fixed time interval, each time using updated short-term predictions of both DRT demand (per zone) and location of idle vehicles. This approach was originally proposed and described in a previous study [6].

One of the most critical elements of providing efficient fleet rebalancing is a proper formulation of the target function. In [6] the target was defined as a linear function of the expected demand $e_{i,j}$, which is the number of trips originating from zone i in a given time horizon (e.g. 30 min) starting from time j:

$$t_{i,j} = a * e_{i,j} + b \tag{1}$$

The linear coefficient a models the relation between the number of incoming requests and the target number of inbound vehicles ready to serve them. The b constant is responsible for setting the desired number of extra vehicles. Both parameters strongly depend on the length of the time horizon and the frequency of fleet rebalancing. In general, they should be relatively low to prevent the rebalancing method from over-reacting which increases empty mileage and reduces availability of vehicles. On the other hand, they should be high enough to actually have a meaningful impact on the actual flow of vehicles. By linking the target value to the expected demand $e_{i,j}$ for the ride-hailing-service, the authors aim to preferably place vehicles in areas of high demand. From an operator's perspective, this should optimize demand throughput and reduce idle time of vehicles while keeping the empty mileage at a reasonable level. However, as explained above, this might lead to high waiting times in peripheral areas with low demand and thus to unequal access to the mobility system.

2.3. Proposed formulations of the rebalancing target

In order to design a ride-hailing service that provides stable waiting times throughout the entire service area, one can modify the calculation of $t_{i,j}$. We decide to investigate these approaches and examine the following formulations:

1. **Estimated demand (ED)**
   This represents computation of the target $t_{i,j}$ according to (1), that is a linear function of the estimated demand for the ride-hailing service, $e_{i,j}$, that originates in zone i from time j within a predefined time horizon. In this study, we use the demand from the previous iteration (day) for estimating $e_{i,j}$, as was proposed in [6].
2. **Fixed number of vehicles (FNV)**
   This is a simple approach to stabilize waiting times where the rebalancing target is set to the same fixed value for all zones. This approach can be referred to as "Supply Anticipation" [29]. However, setting the same target number of inbound vehicles to all non-homogeneous zones that differ in size might lead to an unbalanced distribution of vehicles.
   $$t_{i,j} = b \tag{2}$$
3. **Equal vehicle density (EVD)**
   In this approach, $t_{i,j}$ is defined in a way to obtain an equal number of vehicles per area in all zones, regardless of time j. The target value is defined as
   $$\sum_i t_{i,j} = a * NV * \frac{s_i}{\sum_{k \in Z} s_k} + b \tag{3}$$
   where s_i is the area of zone i and NV represents the current fleet size, which is possibly subject to change over iterations [e.g. if fleet optimization is enabled; 16]. Both FNV and EVD are completely independent from the actually expected demand. Consequently, vehicles might be placed in zones where nobody lives and no demand could possibly be induced.

4. **Equal vehicle-to-population ratio (EVP)**
   This target calculation is similar to the EVD target. The main difference is that the desired vehicle distribution depends on the zonal population. The target value is defined as

$$\sum_i t_{i,j} = a * NV * \frac{p_i}{\sum_{k \in Z} p_k} + b \tag{4}$$

   where p_i is the number of inhabitants in zone i.

Note that if zones have similar population sizes (which is the case for this study), the FNV approach has similarities to EVP. If zones are similarly sized, the FNV approach is very similar to EVD.
This study aims to evaluate the performance of all of these approaches by simulating a scenario at city-scale. Results are compared to the rebalancing strategy proposed by Bischoff and Maciejewski [6], which is represented by ED with a = b = 0.5. This strategy places one vehicle in every zone that experienced one request in the time interval j in the last iteration. For every two additionally experienced requests in j, another vehicle is requested by zone i. In other words, vehicle target numbers have both a demand dependent and a constant component. Corresponding simulations receive an identifier (runId) starting with ED-R. In order to determine the spatial waiting time stability of demand responsive rebalancing, we compare ED with b = 0 as well. Corresponding simulations receive a runId starting with ED. Summarizing, the overall number of rebalancing strategies investigated is five.

2.4. Parameter Variation

In [6], the authors declare that the optimal zonal aggregation level needs to be determined as well as the optimal setting for the rebalancing interval. While they did not investigate different zonal settings, they varied the rebalancing interval length between 15 and 90 minutes for the best setting of a and b and stated no remarkable effects. Dandl et al. [9] investigated the impact of the aggregation level for a demand-anticipatory rebalancing strategy and found that smaller zones lead to better fleet performance in terms of user wait times and empty mileage. The authors state that, for the analyzed Manhattan taxi data set, there is no trade-off between operation costs and service quality in the choice of the zonal aggregation level. However, as this might depend on local properties such as the network topology, we investigate the zonal aggregation level for the use case of Berlin with synthetically produced demand as well. The following parameters are varied for these strategies:

- rebalancing interval length (300 seconds or 1800 seconds),
- zonal aggregation level (447 small LOR zones or 138 large LOR zones or 471 square shaped grid zones),
- the sum of all target values per zone $\sum_i t_{i,j}$ i.e. the agility of the rebalancing systems.

Note that the demand-anticipatory rebalancing approach ED does not allow for direct manipulation of $\sum_i t_{i,j}$. Instead, it only allows for manipulation of the number of vehicles associated with each expected request. With the exception of ED-R, where the parameter set is fixed according to [6], the parameter a and b are chosen such that $\sum_i t_{i,j}$ is close to 1000 vehicles or 2000 vehicles, respectively. In other words: at the moment of rebalancing target computation, the share of vehicles requested by all zones is either 1 or 0.5. Note that FNV requests the same fixed number of vehicles for each zone. Consequently, $\sum_i t_{i,j}$ is a multiple of the number of zones and cannot get set exactly to 1000 vehicles or 2000 vehicles in the investigated use case. Finally, we conduct one single simulation run with rebalancing disabled for comparison (runId base). See the supplemental material [32] for an overview of all runs and their parameter sets as well as the resulting $\sum_i t_{i,j}$ ('-' if not definable) and their results.

2.5. Simulation Setup

To investigate the impacts of the suggested rebalancing strategies, we use a MATSim scenario of Berlin, which is open and freely available [31, also see https://github.com/matsim-scenarios/matsim-berlin] in a development version of the release 5.5. The multi-agent transport simulation MATSim [14] enables a combined simulation of pooled ride-hailing [7] and demand reactions to the service level experienced by the passengers. By its co-evolutionary approach, MATSim allows agents who experience a bad ride-hailing service to move their departure time or choose another transport mode. In comparison to a simulation with a fixed demand this allows for more realism and allows to potentially measure the success of spatially equal wait times in terms of increased demand in previously badly served areas. Moreover, recent developments in the modeling of demand responsive transport systems (DRT) with MATSim are used: Kaddoura et al. [15] developed a module that highly decreases computing times by teleporting DRT users in most iterations and restricts explicit, physical modeling of DRT to a definable share of iterations. In all simulations, DRT users pay 0.20 Euro/km and a base fare of 1 Euro. The DRT system fleet size is fixed to 2000 vehicles. The scenario is run at a 10% sample, i.e. one agent represents ten real inhabitants. Consequently, one ride represents ten rides in reality and one vehicle represents ten vehicles in reality. Alternatively, the results could also be interpreted as a 100% sample of a smaller demand level (which might result from higher prices or lower DRT attractiveness in general). See [17] for a detailed investigation of the effects of different demand sample sizes and demand densities on performance indicators. For the zone rebalancing zone system, we choose "Lebensweltlich orientierte Räume" (LOR), established by the senate of Berlin in 2006 as the official interdisciplinary spatial system for planning, forecasting and investigation [23]. It has three aggregation levels of which the two more detailed ones are used here (447 and 138 zones respectively).

2.6. Performance indicators

The overall aim of this study is to investigate whether the spatial stability of waiting times in the DRT system can be increased while roughly maintaining the profitability of the operator. Let $t_w^{mean}(i)$ be the mean waiting time of all DRT trips originating in zone i. We choose the standard deviation of $t_w^{mean}(i)$ over all $i \in Z$ as the indicator for spatial waiting time equality in this study and denote it as s_W. Winter et al. [29] use a measure called Gini-Index [12], which is commonly used as equality index for income. Both of these measures have issues with varying granularity affecting the comparability. Moreover, we investigate the global mean waiting time, T_w^{mean}, as well as the global 95th percentile of all waiting times, T_w^{p95}, which give an idea of the overall service level provided. While s_W accounts for the spatial dimension of service provision equity, it does not assure service quality. The standard deviation can be low if all agents wait for exactly 2 hours. The 95th percentile of wait times over all rides (not by zone) is used as a measure for the overall service level. In a stochastic simulation of DRT services, it is common that a small number of users experience very high waiting times. This is why we account for 95th percentile instead of the maximum of wait times in order to measure overall service quality. To account for the operator's perspective, we analyze the resulting demand, both on a global and on a zonal level. As rebalancing typically increases the number of empty rides, we investigate the ratio of empty vehicle kilometers travelled (VKT) as well as the overall level of VKT.

A transport policy department interested in spatially equal service provision would need to be interested in a low s_W value. However, to ensure that the operator does not simply lower the quality level in well served area but rather increases quality level in poorly served areas, it needs to regulate the overall level as well. This could be done using the 95th percentile of wait times over all rides. Both performance indicators can be optimized by serving no demand at all. This is why finally the department would have to care about how transport requests are handled. In Germany public transit providers are obliged to accept all passengers with a valid ticket. So, if DRT is to partly replace public transit the DRT operator will not be allowed to reject ride requests that are costly to the operator. In the private sector however, it is not clear how to prevent operators from discarding zones completely in order to minimize s_W or T_w^{p95} by having no rides and thereby no wait time data for the respective zones.

3. Results

In this section, performance indicators of different rebalancing approaches with different parameter sets are analyzed as a first step. Next, a subset of the best simulation runs within each approach is determined and analyzed in more detail. Specifically, the spatial distributions of the mean wait time and the demand are investigated. Moreover, the performance of the selected parameter sets is also tested with the usage of a grid zone system.

3.1. Investigation of rebalancing approaches and parameter sets

All resulting values of the performance indicators can be found in detail for each run in the supplemental material [32]. In the following, we elaborate on the main outcomes of the conducted study. The structure of the result analysis is based on the performance indicators mentioned in 2.6.

3.1.1. Analysis of the demand

Overall, we observe a high variation in the resulting number of rides, with a maximum of 27469 rides and a minimum of 6854 rides. The highest demand values are yielded with small LOR (where number of zones equals 447), both within each rebalancing approach and in general. However, there is a rather strong variation within the zone systems across all rebalancing approaches. For example, the best run with small LOR yields 27469 rides and the worst run with small LOR yields 18006 rides.

Comparing the rebalancing approaches, we observe that EVP achieves a high level of demand in all cases but one. In contrast, the number of rides varies strongly for ED and ED-R (i.e. for demand-anticipatory rebalancing approaches). Thus, the parameter set should be chosen carefully (possibly dependent on the scenario). With FNV, it seems beneficial to put four vehicles into each zone in comparison to a target value of one vehicle per zone. In contrast, for EVD, one can achieve a higher demand by using only half of the idle vehicles for rebalancing.

3.1.2. Analysis of the wait time statistics

Figure 1 suggests a linear correlation between T_w^{mean}, T_w^{p95} and the number of rides. This seems plausible, as the demand reacts to provided service quality. Just like for the demand, one can state a rather high variation in terms of T_w^{mean} and T_w^{p95}. EVP again produces stable results on a high level while small LOR (blue) outperform large LOR (green). Note that the two clear outliers in the top left corner represent runs with demand-anticipatory rebalancing.

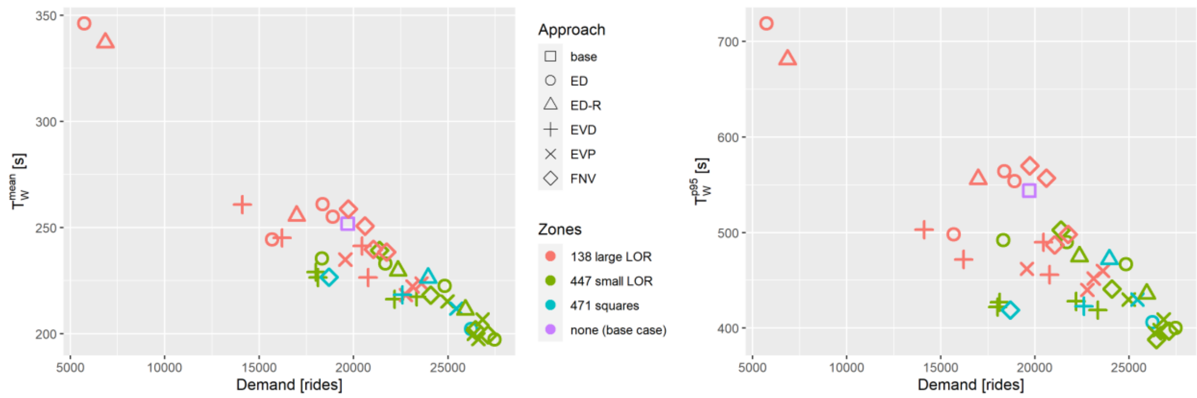

Figure 1 - Average and 95th-Percentile of wait time of all trips in relation to the demand

As shown in the supplemental material [32], the results underline the importance of T_w^{p95} as a performance indicator for the service provision quality. In pairwise comparisons of promising runs one can clearly derive a trade-off between a very low s_W and a very low T_w^{p95}. Consequently, as mentioned in section 2.6, a transport policy department should take this indicator into account, even if service provision equity is a main goal.

3.1.3. Analysis of the empty VKT ratio

Generally, small LOR seem to be correlated with a higher empty VKT ratio. Exceptions seem to be related to low demand which also leads to lower probability for pooling. This could be explained by the effect that for larger zones, the number of vehicles that can stay where they are (at the time of rebalancing target value computation) is higher. In other words, for smaller zones the number of movements needed to satisfy the rebalancing demand is higher. This seems to outweigh the fact that these movements are generally shorter than for large zones.

ED-R and ED produce good and stable results. Placing one vehicle in each zone (FNV) also leads to a low empty ratio. For all other runs, the empty ratio is worse. EVD produces the highest empty VKT ratios, especially when all the idle vehicles potentially get rebalanced (i.e. $a = 1$).

It can be stated that the demand-anticipatory approaches (ED-R and ED) achieve a lower empty VKT ratio in general. This seems plausible as vehicles are only moved if a demand is expected. For example, in late evening hours when the demand level is low, the system is less agile. In contrast, for FNV, EVD and EVP, the system always tries to re-establish its desired state and reacts to every single request that takes a vehicle from one zone to another. Moreover, one can observe that more frequent rebalancing (interval width 300 seconds vs 1800 seconds) leads to a higher empty VKT ratio, which seems plausible. Other than that, the preferred interval width seems to depend on the approach, the zone system and $\sum_i t_{i,j}$ and is not clearly correlated with other outcomes.

3.1.4. Analysis of the service provision equity

In terms of the service provision equity, the small LOR lead to lower s_W values than the large LOR both within each approach and for each parameter set (see also blue symbols versus green symbols in figure 2). EVD achieves best results in terms of service provision equity. Four out of the six lowest s_W values are produced with this approach. However, as figure 2 displays, similar service equality levels in combination with higher demand can be achieved with EVP or FNV (compare blue triangles at the bottom to symbols in the bottom right corner). The worst results are observed for ED-R, FNV and ED when less vehicles are used (i.e. lower $\sum_i t_{i,j}$). Using potentially all idle vehicles with small LOR in FNV and EVP also leads to more equal service provision. This parameter setting yields the best overall results if used with EVD. In general, using a fixed target value for each zone (i.e. FNV, EVD, EVP) seems to lead to higher service provision equity than demand-anticipatory rebalancing. This is in line with the hypothesis formulated in the introduction of this paper.

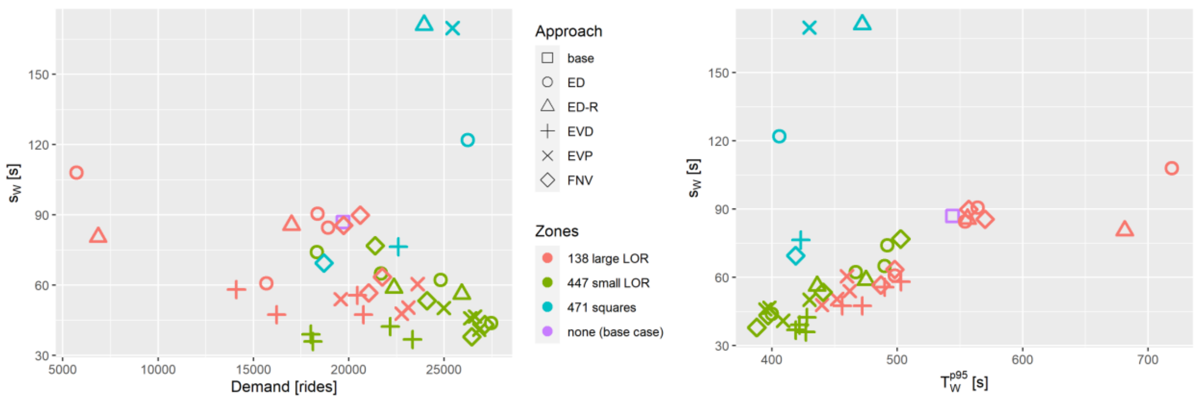

Figure 2 – s_W in relation to the demand and to the 95th percentile of wait time of all trips

3.2. Choice of the best parameter sets and comparison to the usage of a grid zone system

As the aim of this study is to identify rebalancing strategies and parameter sets that lead to more spatially equal mean wait times without significantly affecting the operator's profitability, we focus on a subset of the most promising runs for a more detailed analysis. An operator would be mainly interested in yielding a high demand while keeping the ratio of empty VKT low. The best run of a group is identified by having low s_W and T_w^{p95} values while still maintaining a rather high demand and a rather low ratio of empty VKT. In order to investigate structural differences between the approaches, one parameter set is to be determined for each approach. Please see the supplemental material [32] for a detailed analysis and justification of the chosen set of runs, which consists of ED-R2, FNV6, EVD1, EVP1, ED6.

For the determined subset, we conduct simulations with the exact same parameter sets but replace the LOR zone system with a grid zone system that is composed by squares with a side length of 1.5 kilometers. The size is chosen such that the number of resulting zones is close to the number of small LOR (471 vs. 447). This ensures comparability and that, in FNV, $\sum_i t_{i,j}$ results in roughly the same value. These runs receive the same run identifier as their mirrored runs with an additional '-G'.

In terms of the outcomes, the simulations with a grid zone system perform worse than the mirrored runs with small LOR (see the supplemental material [32]). Concerning the demand, the difference is not very large - with the exception of FNV6 versus FNV6G. The service quality level is considerably lower with a grid system for EVP1, FNV6 and ED-R2. While EVP1-G achieves a considerably lower empty VKT ratio than EVP1, a grid system seems to be associated with a higher empty VKT ratio in any other pairwise comparison. Moreover, the runs with a grid system produce s_W values mostly above 100 seconds whereas most of all other runs yield a s_W value lower than 100 seconds. Figure 2 demonstrates that the grid runs mainly represent clear outliers in the correlation of s_W and the demand or the service quality, respectively. However, the sample of grid runs is considerably smaller than the LOR run sample.

3.3. Detailed graphical comparative spatial analysis of the best simulation runs

For a more detailed analysis, the spatial distribution of waiting times and demand is visualized and compared to the base case. In order to validate the runs in terms of the service provision equity, the mean wait times per zone are analyzed first. In the following, the resulting number of rides per zone is evaluated in order to determine changes in the spatial structure (e.g. more people in the outskirts use the ride-hailing service).

3.3.1. Analysis of the spatial distribution of the mean wait time

Figure 3 shows the mean wait time for the ride-hailing service per rebalancing zone for the runs described in 3.2. As initially expected, the demand-responsive standard setup (ED-R2), proposed in [6], serves the outskirts with a considerably lower service quality than the city center. However, the ED6 setup demonstrates that demand-responsive rebalancing can also lead to spatially stable mean wait times. Note that this run yields the best overall mean wait time and the highest demand.

Another interesting aspect is that the spatially more equal wait times of runs EVD1 and EVP1 in comparison to the reference rebalancing approach ED-R2 are due to both the outer zones being served at lower average wait times and the inner zones being served at slightly higher wait times. That means EVD1 and EVP1 redistribute service levels in a way which can help to render DRT more of a complement to public transit than the reference rebalancing strategy ED-R2 does, i.e. they provide better service where public transit is less frequent in the outskirts of the city and slightly worse service in the city center which is well served by existing public transit. ED6 instead seems to have almost identical mean wait times in the city center as ED-R2, but has significantly lower wait time in the outer zones. This is also reflected by the lower average wait time over all rides [32].

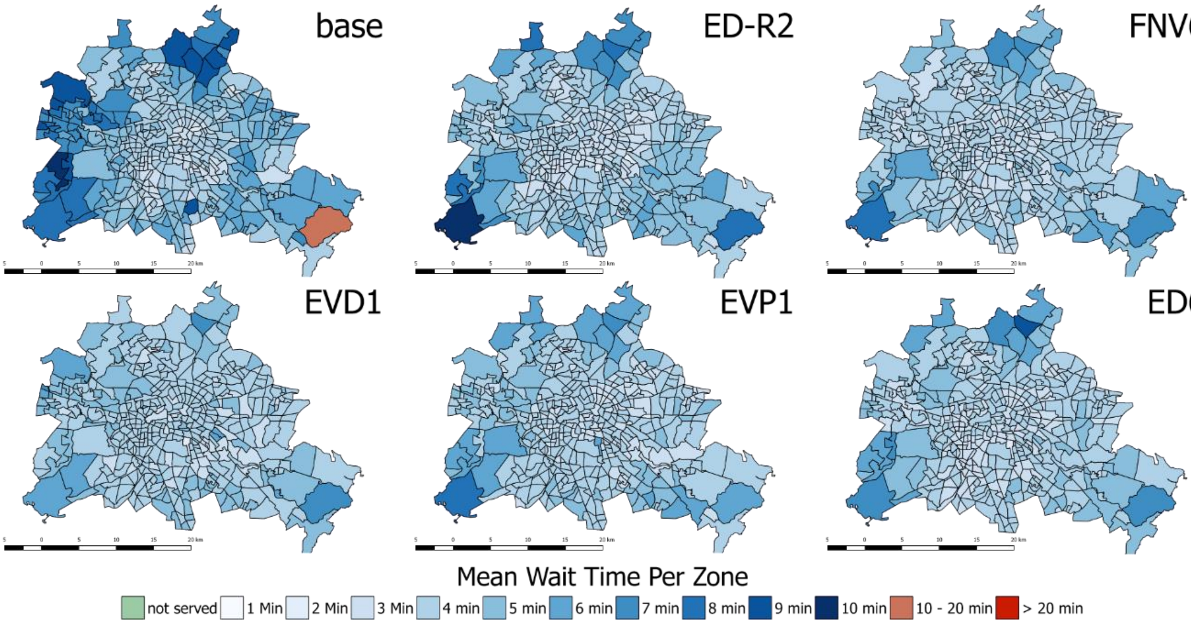

Figure 3 – Spatial Distribution of the Mean Wait Time

3.3.2. Analysis of the spatial distribution of the demand

In the introduction, the concern was stated that purely demand-responsive rebalancing could lead to significantly less demand in the outskirts. As stated in section 3.3.1, outskirts are served in acceptable manner with ED6. Therefore, it is important to look on the spatial distribution of the demand as well and assure that the acceptable mean wait times in the outskirts of ED6 do not correspond to very few lucky requests. Generally, we aim to investigate whether the improved equality of the service level has an impact on the spatial demand distribution.

Figure 4 displays the spatial demand distribution. In contrast to the stated concerns, ED6 does not yield a lower demand level in the outskirts compared to the other runs. Overall, the demand reaction takes place in the inner city and the southern outskirt zones. Apart from the central south however, the number of rides is more or less constant through the approaches. The demand does not react to the improved service quality in large border zones (compare the spatially more equal FNV6, EVD1, EVP1 and ED6 with the reference rebalancing ED-R2 in figure 3). This can be explained by the way the service is designed. Through all of these simulations, the mean distance between customer's start and end point varies from 3211 meters to 3231 meters. While the service operator charges a base fare of 1 Euro, every kilometre travelled costs 0,20 Euro. This has the consequence, that the service is attractive for a specific distance range only and long distances are avoided. As the activity and population density in the large LOR is rather low by definition [23], trips starting there require rather large distance transport. That is why, for most agents originating or driving to these zones, the ride-hailing service is no attractive option. This could be tackled with another fare system but does not seem to be related to the mean wait time in this specific case study.

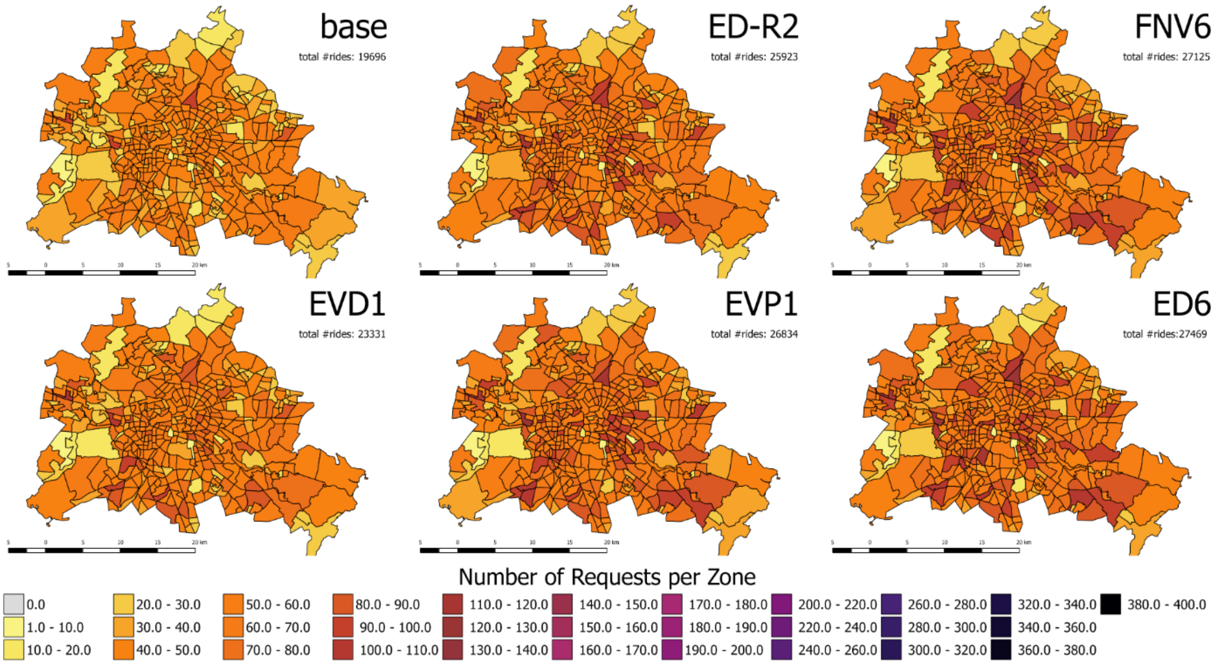

Figure 4 – Spatial Demand Distribution

Note the non-linear colour scale that also covers the plot in the supplemental material [32] for comparability.

3.3.3. Detailed comparative analysis to the usage of grid zones

In the previous subsections we have seen that large zones at the outer range of the city tend to be served with lower quality and induce less demand than central zones. In order to evaluate whether this is solely an artefact of the zone size and shape, we compare the spatial distribution of the mean wait time and the demand to the usage of a grid zone system. This means all zones have the same size. Please see the supplemental material [32] for the detailed analysis including the plots. See also section 3.2. The results of the comparative analysis indicate that the spatial structure of both the demand and the mean wait times does not change very much when switching from LOR to square zones. However, the level of the demand is considerably lower for a grid of square zones and the number of rides is equal to 0 in most of the border zones.

4. Discussion

While an equal service level could bring an element of social justice, it can be argued that it could have a negative impact on land use development and possibly lead to urban sprawl. Usually, public transit network density (and thus waiting times) also differ within large city areas. However, the purpose of this study is to investigate methods that prevent the competition of ride-hailing services in city centers that are already sufficiently served by public transport and make those services attractive in outside areas, where the transport system could benefit from a complementary mode.

Dandl and Bogenberger [8] found that the zonal aggregation level is decisive for rebalancing performance and state that smaller zones work better. In this study, we observed the same outcomes by comparing large and small LOR zone systems. Additionally, we investigated the impact of the zone shape and the demand reaction to it. Depending on the rebalancing strategy, the shape of zones can have an impact on both the service quality (and thus the demand) and the profitability (i.e. empty VKT ratio). The results suggest that with smaller rebalancing zones, more people tend to use the ride-hailing service.

Winter et al. [29] investigated rebalancing in the context of parking constraints. In their study, the aspect of pooling is not considered, which is incorporated here. This paper underlines their findings that demand-anticipatory rebalancing is generally associated with higher service inequality compared to supply-anticipatory rebalancing. Furthermore, the authors determined a correlation between service equality and mean wait times. We observe the same tendency, with the exception of the grid runs.

5. Conclusion

We show that in the case study of Berlin, DRT users experience different service quality levels depending on where their trip starts. This can be seen as a social inequality or imbalance in the ride-hailing system. We investigate five different rebalancing approaches to minimize standard deviation of mean waiting times per zone under the consideration of mode choice behavior. In terms of service provision equity, it seems to be a good solution to maintain equal vehicle density by rebalancing throughout the ride-hailing service area. The distribution of mean wait times per zone can be narrowed. Demand rises in consequence to improved service provision equity with the exception of areas where the pricing scheme is not attractive enough. However, resulting demand, wait time statistics as well as the empty VKT ratio could turn out significantly different for all of the investigated rebalancing strategies when applied in rural areas. This should be addressed in future research on this topic. This study underlines the findings of other studies which state that large rebalancing zones lead to lower system efficiency and lower service quality. Additionally, we observe a reduction of the demand due to large rebalancing zones.

Acknowledgement

This work was partly funded by the German Federal Ministry of Transport and Digital Infrastructure (funding number 16AVF214).

References

1. Albert, M., Ruch, C., Frazzoli, E., (2019). Imbalance in mobility-on-demand systems: A stochastic model and distributed control approach, *ACM Transactions on Spatial Algorithms and Systems (TSAS)* 5, 1–22.
2. Alonso-Mora, J., Samaranayake, S., Wallar, A., Frazzoli, E., Rus, D., (2017). On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment, *Proceedings of the National Academy of Sciences 114, 462–467,* doi:10.1073/pnas.1611675114.
3. Babicheva, T., Burghout, W., Andreasson, I., Faul, N., (2018). The matching problem of empty vehicle redistribution in autonomous taxi systems. *Procedia computer science 130*, 119–125.
4. Bischoff, J., Maciejewski, M., (2016a). Autonomous taxicabs in Berlin – a spatiotemporal analysis of service performance, *Transportation Research Procedia 19*, 176–186. doi:10.1016/j.trpro.2016.12.078.
5. Bischoff, J., Maciejewski, M., (2016b). Simulation of City-wide Replacement of Private Cars with Autonomous Taxis in Berlin, *Procedia Computer Science 83*, 237–244. doi:10.1016/j.procs.2016.04.121.
6. Bischoff, J., Maciejewski, M., 2020. Proactive empty vehicle rebalancing for demand responsive transport services, *Procedia Computer Science 170*, 739–744. doi:10.1016/j.procs.2020.03.162.
7. Bischoff, J., Maciejewski, M., Nagel, K., (2017). City-wide shared taxis: A simulation study in Berlin, in: *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, IEEE. doi:10.1109/itsc.2017.8317926.
8. Dandl, F., Bogenberger, K., (2018). Comparing future autonomous electric taxis with an existing free-floating carsharing system, *IEEE Transactions on Intelligent Transportation Systems 20,* 2037–2047.
9. Dandl, F., Hyland, M., Bogenberger, K., Mahmassani, H.S., (2019). Evaluating the impact of spatio-temporal demand forecast aggregation on the operational performance of shared autonomous mobility fleets, *Transportation 46*, 1975–1996.
10. Fahnenschreiber, S., Gundling, F., Keyhani, M.H., Schnee, M., (2016). A multi-modal routing approach combining dynamic ride-sharing and¨ public transport, *Transportation Research Procedia 13*, 176–183.
11. Ford Jr, L.R., Fulkerson, D.R., 1956. Solving the transportation problem, *Management Science 3,* 24–32.
12. Gini, C., 1912. Variabilita e mutabilita: contributo allo studio delle relazioni statistiche, *Studi Economico-giurdici,* Facolta di Giurisprudenza della R. Universita di Cagliari 3.

13. Hörl, S., Ruch, C., Becker, F., Frazzoli, E., Axhausen, K.W., (2017). Fleet control algorithms for automated mobility: A simulation assessment¨ for Zurich. Annual Meeting Preprint 18-02171, *Transportation Research Board. Washington, D.C.*
14. Horni, A., Nagel, K., Axhausen, K.W. (Eds.), (2016). The Multi-Agent Transport Simulation MATSim. *Ubiquity*, London. doi:10.5334/baw.
15. Kaddoura, I., Leich, G., Nagel, K., (2020a). The impact of pricing and service area design on the modal shift towards demand responsive transit, *Procedia Computer Science 170*, 807–812. doi:10.1016/j.procs.2020.03.152.
16. Kaddoura, I., Leich, G., Neumann, A., Nagel, K., (2020b). A Simulation-based heuristic for the improvement on-demand mobility services. VSP Working Paper. *TU Berlin*, Transport Systems Planning and Transport Telematics.
17. Kaddoura, I., Schlenther, T., (2020). The impact of trip density on the fleet size and pooling rate of ride-hailing services: A simulation study. VSP Working Paper 20-36. *TU Berlin*, Transport Systems Planning and Transport Telematics.
18. Krueger, R., Rashidi, T.H., Rose, J.M., (2016). Preferences for shared autonomous vehicles. *Transportation Research Part C: Emerging Technologies 69*, 343–355. doi:10.1016/j.trc.2016.06.015.
19. Maciejewski, M., Bischoff, J., (2018). Congestion effects of Autonomous Taxi fleets, *Transport 33*, 971–980. doi:10.3846/16484142.2017.1347827.
20. Mulley, C., Nelson, J.D., 2009. Flexible transport services: A new market opportunity for public transport, *Research in Transportation Economics 25*, 39–45.
21. Nagel, K., Bischoff, J., Leich, G., Maciejewski, M., (2018). Simulationsbasierte Analyse der Wirkungen von Flotten autonomer Fahrzeuge auf städtischen Verkehr, *Zeitschrift für Verkehrswissenschaft*, 197–224.
22. Navidi, Z., Nagel, K., Winter, S., (2019). Toward identifying the critical mass in spatial two-sided markets, *Environment and Planning B: Urban Analytics and City Science,* 239980831984218doi:10.1177/2399808319842181.
23. Senatsverwaltung für Stadtentwicklung und Wohnen Berlin, 2020. Lebensweltlich orientierte Räume (LOR) in Berlin: Planungsgrundlagen. URL:¨ https://www.stadtentwicklung.berlin.de/planen/basisdaten_stadtentwicklung/lor/.
24. Shaheen, S., (2018). Shared Mobility: The Potential of Ridehailing and Pooling, *Island Press/Center for Resource Economics*, Washington, DC. pp. 55–76. URL: https://doi.org/10.5822/978-1-61091-906-7_3, doi:10.5822/978-1-61091-906-7_3.

25. Spieser, K., Samaranayake, S., Gruel, W., Frazzoli, E., (2016). Shared-vehicle mobility-on-demand systems: A fleet operator's guide to rebalancing empty vehicles, in: Transportation Research Board 95th Annual Meeting, *Transportation Research Board*.
26. Stiglic, M., Agatz, N., Savelsbergh, M., Gradisar, M., (2018). Enhancing urban mobility: Integrating ride-sharing and public transit, *Computers & Operations Research 90*, 12–21.
27. Syed, A.A., Akhnoukh, K., Kaltenhaeuser, B., Bogenberger, K., (2019). Neural network based large neighborhood search algorithm for ride hailing services, in: EPIA Conference on Artificial Intelligence, *Springer*. pp. 584–595.
28. Wasa, Y., Tanaka, T., (2019). Path-integral control for optimal vehicle routing over nonlinear and dynamic traffic networks, in: *2019 18th European Control Conference (ECC)*, pp. 2189–2196. doi:10.23919/ECC.2019.8796248.
29. Winter, K., Cats, O., Martens, K., Arem, B., et al., 2020. Relocating shared automated vehicles under parking constraints: assessing the impact of different strategies for on-street parking, *Transportation*, 1–35.
30. Zhang, Y., Zhang, Y., (2018). Exploring the relationship between ridesharing and public transit use in the united states, *International journal of environmental research and public health 15*, 1763.
31. Ziemke, D., Kaddoura, I., Nagel, K., (2019). The MATSim Open Berlin Scenario: A multimodal agent-based transport simulation scenario based on synthetic demand modeling and open data, *Procedia Computer Science 151*, 870–877. doi:10.1016/j.procs.2019.04.120.
32. Schlenther, T., Leich, G., Maciejewski, M., Nagel, K., (2021). Supplemental Material to Addressing Spatial Service Provision Equity for Pooled Ride-Hailing Services through Rebalancing. *TU Berlin*, Transport Systems Planning and Transport Telematics. doi: http://dx.doi.org/10.14279/depositonce-11337.