SIMULATION-BASED OPTIMIZATION OF SERVICE AREAS FOR POOLED RIDE-HAILING OPERATORS

Joschka Bischoff (corresponding author)
Berlin Institute of Technology (Technische Universität Berlin)
Transport Systems Planning and Transport Telematics
Salzufer 17-19
10587 Berlin
Germany
Telephone: +49 30 314 29521
FAX: +49 30 314 26269
bischoffvsp.tu-berlin.de

Michal Maciejewski
Poznan University of Technology
Division of Transport Systems
ul. Piotrowo 3
60-965 Poznan
Poland
Telephone: +49 48 61 647 5957
michal.maciejewski@put.poznan.pl

Ihab Kaddoura
Berlin Institute of Technology (Technische Universität Berlin)
Transport Systems Planning and Transport Telematics
Salzufer 17-19
10587 Berlin
Germany
Telephone: +49 30 314 78793
kaddoura@vsp.tu-berlin.de

Kai Nagel
Berlin Institute of Technology (Technische Universität Berlin)
Transport Systems Planning and Transport Telematics
Salzufer 17-19
10587 Berlin
Germany
Telephone: +49 30 314 23308
nagel@vsp.tu-berlin.de

Submission date: December 18, 2017
4,375 words + 5 figures + 1 table = 5,875 words

Abstract

Dynamic ride hailing with passenger pooling has become a popular form of urban transport and is a growing sector around the globe. The area where these services operate is often limited to densely populated inner city districts, whereas non-pooled options are often available in larger areas. In this paper, we introduce a simulation-based methodology that allows to optimize the service area of a ride hailing service using an agent-based simulation and apply it to the taxi demand of Berlin, Germany. Three different criteria are used for the optimization, which take the average vehicle occupancy, the revenues collected per area or both into account. The results show that for the given parameters a service area that focuses on an extended central area and some areas around may be profit-maximizing for operators.

INTRODUCTION

Dynamic ride-hailing (sometimes also referred to as ride-sharing) services have been growing world-wide in the last years, disrupting taxi markets and both the private and public transport sector all over the globe. More recently, ride-sharing operators, or transportation network companies (TNCs), have started pooling customers with similar headings (1, 2). On one hand these pooled services offer generally lower fares to customers, who may have to take accept a certain detour and discomfort. On the other hand, pooling may reduce the overall vehicle miles traveled (VMT) and thus mitigate negative congestion and environmental effects arising from additional empty mileage of ride-sharing vehicles. With possible future fleets of shared autonomous vehicles (SAVs) offering taxi like services, the additional VMT may lead to a substantial increase in congestion (3) and pooling rides may be one way out of this.

Generally, (profit-oriented) TNCs offer their pooled services in smaller, often more densely populated areas than their non-pooled services. The availability of pooled services in an area may also depend on the time of day, vehicle availability or other factors. In this paper, we propose an agent-based simulation approach to determine in which areas offering pooled rides is economically justified. The approach is applied to a real-world demand in Berlin, Germany.

STATE OF THE ART

While pooled rides in dynamic ride-hailing applications are still a relatively new idea, the general concept of sharing a taxi or limousine with a third party is older. Traditionally, these ran along (somewhat flexible) routes, which allows for an easy matching of passengers. Examples for such systems exist, e.g, in Santiago de Chile, South Africa, several post-soviet countries and the Arabic world (4, ch. 2). In the United States, the majority of these jitney services were banned in the 1910's, mainly because they were seen as an unwanted competition to streetcar operators. Since then, shared rides have been mainly used for special user groups (such as para-transit for the disabled) or as shuttle services to connect, e.g., with airports. Somewhat related to them, demand-responsive transport systems (DRT) have filled a niche in more rural areas, especially in Europe. Dynamic pooled ride-hailing could only develop thanks to an increased connectivity and smart-phone usage both on the customer's and dispatcher's side.

In science, pooling in dynamic ride-hailing has been looked at mainly from the algorithmic side of matching customers. This problem is usually dealt with by means of various heuristic methods (5, 6, 7) that are typically paired with a transport simulation software. Other questions, such as the willingness of customers to accept pooled rides and the acceptance of detours have been discussed mainly in the context of car-pooling, rather than ride-hailing. Some results are likely to be similar, such as the general acceptance for detours or the perception of travel time, while constraints to timing and loss of flexibility (8) would not apply to ride-hailing.

A methodology to evaluate service areas and their importance to operators using geographically weighted regression models was presented for free-floating car-sharing services in Berlin (9), essentially predicting profitable areas based on historical data and areas of interest. In a previous simulation study focusing on the possible impact of (non)-pooled SAVs, we have found a significantly higher share of empty mileage in areas with low demand (10).

In real-world applications, TNCs have been constantly adapting their pooled services, partly modifying them in a field study like style in some areas at first and rolling it out on a larger scale afterwards. For example, when Uber started offering pooled services, the financial risk for a match was on the customer's side, who only received a significant discount on his fare if a second

customer boarded the ride. Since then, this has been transformed to a flat-fare system with a guaranteed arrival time window, where the match-finding risk is on the operator's side and prices are dynamically adjusted depending on the likelihood of finding other customers along the route (11). Furthermore, special discounts or flat fares may be offered for travel within some high-density zones, where the likelihood of pooling is overall higher (12). Lyft is operating according to similar procedures. This makes pooled TNCs a direct opponent of public transport in many markets(13), though in some cities co-operations between both have evolved.

A large service area paired with a small fleet was one of the reasons leading to the economic failure of *Kutsuplus*, a DRT system in Helsinki, Finland (14), which was canceled after less than two years of service after a well-observed start. This stresses out the importance of choosing an economically sustainable business area.

METHODOLOGY
In this study, we use an iterative, simulation-based approach to evaluate the potential of zones in a service area for a dynamic, pooled ride-hailing service.

Simulation framework
We chose MATSim (Multi-Agent Transport Simulation) (15) as the simulation software to work with. MATSim is an activity-based and dynamic transport model. It is available as open source software and written in JAVA and offers many plug-in points for an easy extension of its functionality.

Transport simulation
The demand for transport is modeled with individual agents. Each agent holds one or more plans which describe the (often) daily activity schedule as well as the travel in between activities by different transport modes. Initial plans have to be provided and these may be modified during the process of demand adaptation to supply. The demand adaptation is based on an evolutionary iterative approach in three steps: (1) travel plans are executed (mobility simulation with a mesoscopic queue-based traffic flow model), (2) the executed plans are scored (evaluation) and (3) plans are modified (learning). This process is repeated for several iterations and typically results in an approximate stochastic user equilibrium (*16*). A detailed description of the simulation framework is provided in Raney and Nagel (*17*). The iterative process is also depicted in the center of Fig. 1.

Simulation of dynamic pooled ride-hailing
The simulation of dynamic transport modes, such as taxi, DRT or SAVs is based on MATSim's extension for dynamic vehicle routing problems (DVRP) (18, 19). It consists of a framework for an on-line dispatch of one or several fleets of vehicles while the mobility simulation is running. In this paper, we are using the DRT extension, which may be used for both pooled taxi rides or typical demand responsive transport use cases. When a ride request reaches the dispatcher, it is assigned to such a vehicle that an increase in the overall time-wise detour is lowest. This happens under the condition that (1) the travel times for the passengers currently in the vehicle or awaiting it and the new customer do not increase beyond specified thresholds and that (2) the expected boarding times for the awaiting customers and the new one are within a requested time frame. Should no suitable vehicle be available, or the request be deemed invalid because its start or end location is outside the service area, the request is rejected. A more detailed description of the extension and
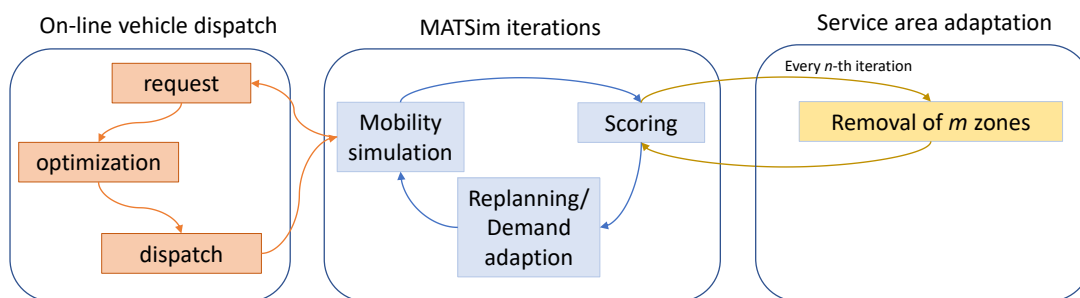
FIGURE 1 Schematic overview of the simulation cycles, extended from (18)

its underlying algorithm is available (*5*). In Fig. 1, the role of DRT extension is depicted on the left side.

Service area adaption

In order to provide adaptability on the operator's side, the existing integration of the multi-iterative simulation in MATSim and the on-line vehicle dispatch provided by DVRP has been extended by a third component that adjusts the service area a ride-hailing operator serves. To achieve this, an initial service area is defined by a set of similarly sized shapes. For each zone, several spatial optimization criteria are collected within each iteration of a MATSim run. After each *n*-th iteration, these criteria are evaluated and a certain number *m* of zones performing worst are removed. The new, smaller service area is then used for the next *n* iterations, and the process of evaluation and adaptation is repeated (as depicted on the right side in Fig. 1). Depending on the fleet used and the spatial distribution of requests, this will, after a certain number of iterations, lead to a service area where, e.g., the operator profit can be maximized. Since the fleet size is kept fixed throughout all iterations, a continuation of the adaptation process will further reduce the service area, but lead to a decrease in profits, as the fleet is too big. There is no adaption on the demand side between iterations.

Optimization criteria

In this study, we use three different optimization criteria:

1. The contribution of each zone to the operator's total revenue. This can be defined as the half of the sum of all revenues generated by all trips starting or ending in zone i (trips starting and ending in zone i are counted twice into the same zone).

$$R_i \tag{1}$$

2. The average occupancy of all vehicles passing through a zone, defined by the person miles traveled in a zone i (PMT_i)[1] and the overall vehicle miles traveled within that

---

[1]The driver of the vehicle does not count as a person. This means, an empty driving vehicle has an occupancy of zero.
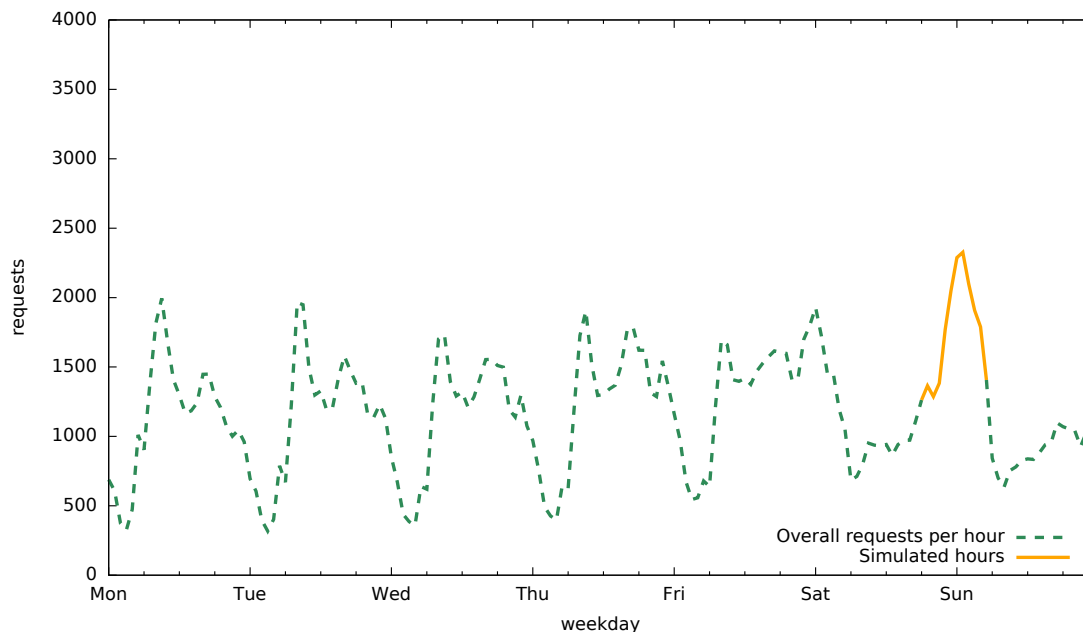
FIGURE 2 Weekly demand for taxi rides observed

zone (VMT_i).

$$\rho_i = \frac{PMT_i}{VMT_i} \tag{2}$$

3. The multiplication of 1. and 2.

$$P_i = \rho_i \cdot R_i \tag{3}$$

A sole optimization based on revenue will not take operator costs into account and may lead to zones being served which attract long trips. These may mostly be non-pooled (due to the distance traveled) and a vehicle may need to drive empty for a long time afterwards to reach the next customer. Both aspects may or may not be part of the fares charged and thus the revenue collection. An optimization based on the average zonal vehicle occupancy assumes that a higher vehicle occupancy generates a higher overall profit. This, once more, depends on the fares charged to customers. A combination of both parameters may balance somewhat in between both.

DEMAND ESTIMATION
The demand used in this paper is based on GPS trajectories and status messages collected by the biggest radio taxi dispatch center in Berlin, Germany. The dataset covers roughly half of the city's fleet of 8,000 vehicles and covers several weeks in spring 2013 and 2014. During one week, more than 200,000 rides were registered (see Fig. 2). The average trip distance is around 7 km, or 4.3 mi. From Monday to Friday, there is a considerable morning peak around 8 am and an afternoon peak which is followed by a strong evening demand on Fridays. The weekly absolute demand peak can be observed during Saturday nights, while there is less demand on Sundays. This demand pattern
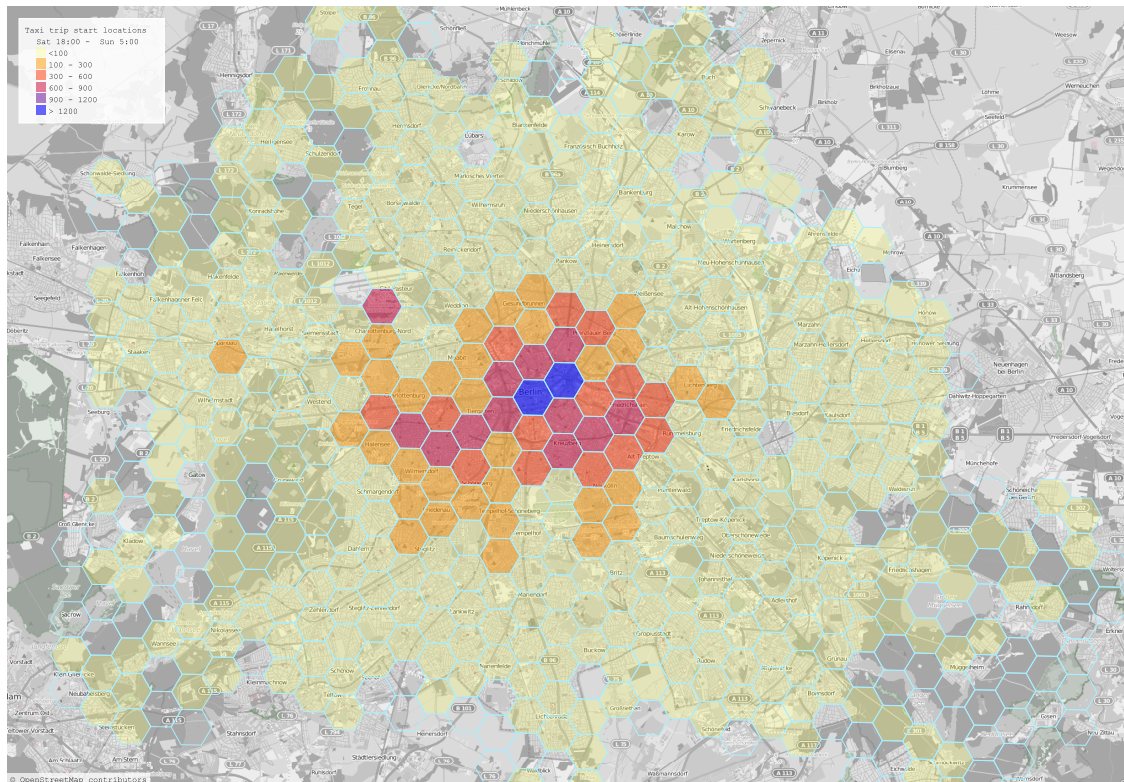
FIGURE 3 Taxi trip start locations during a Saturday night

for taxi rides is qualitatively similar to the one observed in other cities, including New York City (*20*). A detailed description of the taxi market, including a spatial analysis, is also available (*21*).

The overall taxi market in Berlin is strongly fragmented – the majority of taxi companies owns only one or two vehicles and there are only few companies with a fleet of 50 or more vehicles. While there are several radio dispatch agencies, each operator is working for himself and there is no centralized dispatch optimization.

In 2017, there are three ride hailing operators active in Berlin, but the legislative restrictions are high (e.g., fleets are often owned by the operator rather than the drivers) and the fleets, service areas and operation times in use are small, so taxis remain the pre-dominant provider. However, legislation is undergoing changes, that may influence the whole sector.

The focus in this paper lies on the demand peak during Saturday nights between 6 pm and 5 am on Sunday morning (marked in orange in Fig. 2), because this is the time where the willingness to pool rides among taxi users is likely to be highest as many trips are rather leisure-oriented. Demand during this time of week is generally originating from city center locations (and also from Tegel Airport until its closure at 11 pm). Fig. 3 provides an overview of the demand distribution. Since the trajectory data covers only half of the fleet, the overall demand used for simulation purposes is scaled up accordingly. With the focus of this study being a service area optimization for one single operator, we assume that a random 10 % of all requests, or around 3,000 trips in absolute numbers, per iteration are submitted to and potentially served by this operator.

SIMULATION SETUP
Supply
A fleet of 50 vehicles with a capacity of 3 is used in our simulations. These are initially distributed throughout at popular taxi locations within the city center. If a vehicle is unassigned, it does not return to a rank but remains idle at its last location until a the vehicle is dispatched again. Vehicles are available from 6 pm to 5 am the following day. Breaks of drivers are not modeled. Operator costs are assumed to be a fixed 150 EUR per day and vehicle. Vehicle costs were assumed to be ~50 EUR and driver costs ~100 EUR. The additional cost per km is assumed to be 0.30 EUR.

Fare structure and operator profit
We assume a fixed fare based on the direct (unpooled) distance of a trip and assume a fare of 50 % of a regular taxi trip. This means a base fare of 1.95 EUR, and a de-grading distance fare of 1 EUR per km for the first 7 km and 0.75 per km thereafter. The operator profit Π can be calculated by summing up all revenues and deducting the sum of costs.

Dispatch
A vehicle may be dispatched to a request, if the following thresholds are kept for both the new request and all passengers currently on board or awaiting the vehicle: (a) The overall travel time detour of the trip, including the initial wait time, is below 50 % plus seven minute (to allow a pooling also on short trips). (b) The expected maximum wait time is below seven minutes. These parameters were found to be well-working in densely populated areas (5) and warrant a good service quality.

Zonal system and optimization
Initially, we place a grid of 567 hexagons with a radius of 1.5 km (or 0.93 mi) over Berlin. After iteration 10, all zones are removed where not a single trip has started or ended in the preceding iterations. This removes, e.g. water, forest or industrial areas. For the optimization process, as described above, $n = 10$ and $m = 10$ was set, i.e., after each 10th iteration 10 worst performing zones were removed. This continues as long as there are more than $2 \cdot n = 20$ zones left to remove and ends after 360 iterations.

RESULTS
For the chosen set of parameters and fleet, we evaluate the average operator profit per each *n* iterations (Π^{mean}). The highest operator profits are experienced when using the factor P_i as an optimization criterion, though the differences to occupancy (ρ_i) or zonal revenue (R_i) are slim.

Operator Revenues and Profit
The initial operator profit without any service area restrictions is around 395 EUR for the whole fleet. Removal of zones for all optimization criteria leads to an increased Π^{mean}, as Fig. 4 reveals. A pure optimization based on revenues per zone performs somewhat worse than the other two criteria. Π^{mean} is the highest for iterations 240–250, where it is the *P*-oriented optimization with 1,698 EUR and 1,615 EUR for the occupancy-based ρ-optimization. For the revenue-based (i.e. *R*-oriented) optimization, it is the highest in iterations 260–270 with 1,508 EUR. A further removal of zones leads to a reduction in operator profit due to an over-dimensioned fleet. From iteration 310 (for ρ) or 330 (for *R* and *P*), the operator starts loosing money. Table 1 provides an overview
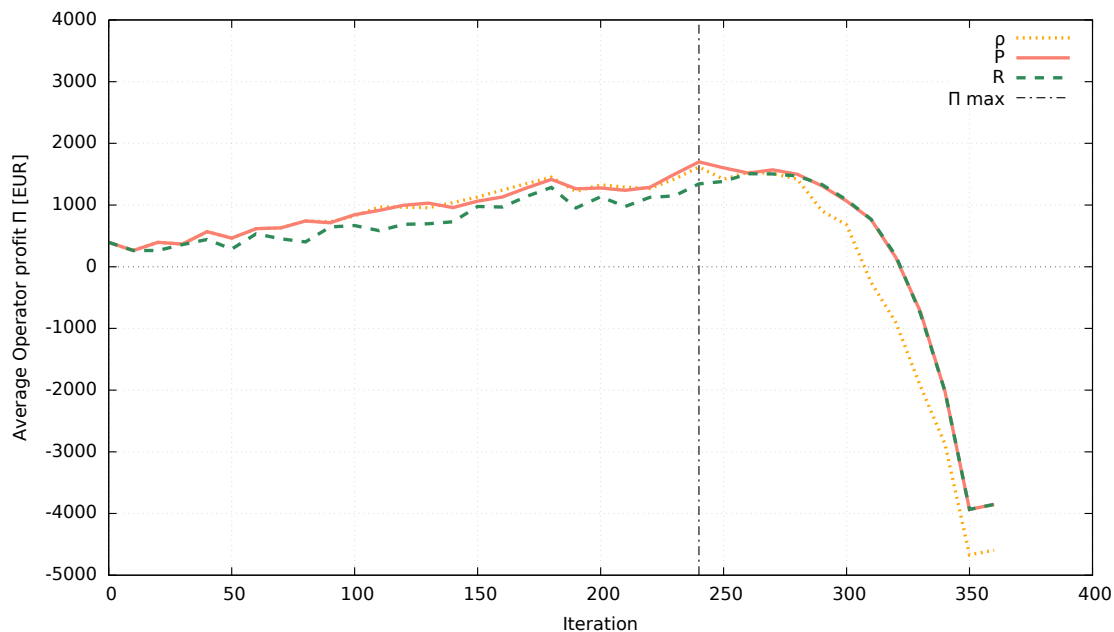
FIGURE 4 Average operator profit per n-iterations

TABLE 1 Average performance indicators for different simulation setups

Criterion	Iterations	Trips	Revenue	Cost	Overall Distance	Empty Distance
			[EUR]	[EUR]	[km / vehicle]	[km / vehicle]
R	260-270	1808	12 479	10 971	231	49
ρ	240-250	1820	12 656	11 041	236	50
P	240-250	1840	12 773	11 075	238	50

of the average performance indicators for all three optimization criteria for those iterations with the highest profit. Both revenue, absolute number of trips and the cost incurred are the highest for trips using the ρ-optimization of the service area. Also the vehicle utilization is the highest, with an average of 238 km per vehicle and shift. This value seems feasible, though it is more than what ordinary taxis drive in the same time (due to a general oversupply of taxis in Berlin). The values for the simulation runs using the other optimization criteria are somewhat similar, with the values for the revenue-based R-optimization being constantly below those of the occupancy-based ρ-optimization of the service area.

Service area
Fig. 5 shows the resulting service area (green) for using each of the optimization criteria. In each case, the city center, defined as the inner circle overground railway, is covered in full. Also the area southwest of the city center (districts of *Wilmersdorf*, *Steglitz* and parts of *Zehlendorf*) are covered in all cases. For the revenue-based optimization (top left in the figure), only the core area in the center is contiguous, while there are some unserved areas directly around it. Three outlying zones to the west, east and south mark spots where there is a certain amount of revenue, which may be a result of longer trips ending in these residential areas.
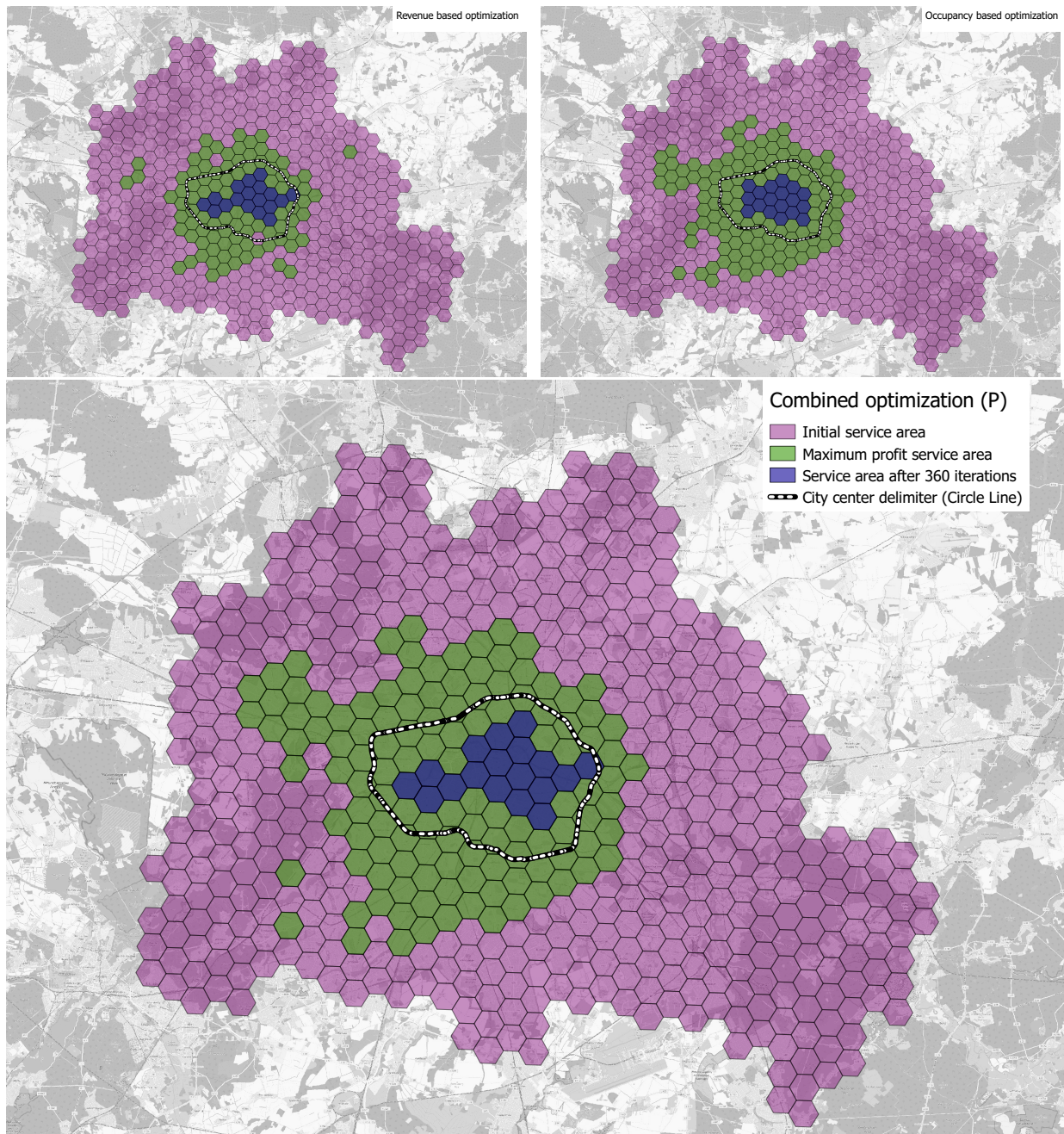
FIGURE 5 Resulting service area for revenue-based (top left), occupancy-based (top right) and performance-based optimization.

Using the optimization approach based on occupancy, the resulting area is contiguous (with one exception). The service area now also covers the link leading to *Spandau*, a well-populated sub-center in the west of the area. Further outlying areas are not covered, indicating that vehicles traveling there are not usually occupied by more than one person.

Using the factor of both in the P-optimization, the resulting area is overall very similar to the occupancy-based approach. However, the area is cut somewhat differently especially in the

west and southwest. This may, however, result from stochastic effects.

A look at the service area remaining in the last iterations (marked in blue), however, reveals a similar service area for both the P-based and revenue-based optimization approach. This area contains both the eastern and western city center cores around *Alexanderplatz* and *Zoologischer Garten* as well as classical nightlife hot spots in the south east of the center. The occupancy-based approach results cuts the zones differently and the resulting areas form a more compact area. This is generating less profit, as Fig. 4 reveals for the last iterations.

Quite notably, the maximum profit service area overlaps to a large extent with the one where other operators of mobility-on-demand services, such as Free-Floating car sharing and bike sharing companies, offer their service.

CONCLUSION

In this paper, we were able to demonstrate a methodology to optimize the business area for a dynamic and pooled ride-hailing service. The iterative approach allows a specific and modifiable definition of service and optimization criteria. The selected use case for the city of Berlin is of special interest, as ride-hailing is still heavily discussed in Germany and an efficient pooling of passengers may increase the overall efficiency of city traffic and thus the acceptance of these services as a whole. In the illustrative use case described, the profit of a ride-hailing operator operating with a fixed fleet size during the week's busiest time, may be maximized by focusing its service area in an extended area around the city center. In this case, the choice of an optimization criterion that takes both revenues and the occupancy of vehicles into account scores maximizes profits.

The use case presented here has some limitations. Namely, prices for pooled ride-hailing options are usually very dynamic and take into account the likelihood of matching a customer with another given the route and time of day. This price prediction influences both the ridership (who may opt to use a different mode instead) and the profitabilies of zones and should be taken into account at a later stage. Furthermore, modeling the choice between pooled and non-pooled options is another interesting field that could be looked at. Finally, it may be interesting to use the optimization framework in the context of future fleets of shared and pooled autonomous vehicles. In this field, an optimization based on the system welfare that takes the user benefits into account may provide an insight on where an operation may not be profitable for operators, but socially beneficial and should possibly be supported in some way.

References

[1] Uber, uberPOOL – Sharing is saving, 2017,
`https://www.uber.com/nyc-riders/products/uberpool/`, accessed 12.05.2017.

[2] Lyft, *Meet Lyft Line*, 2017,
`https://www.lyft.com/line`, accessed 17.07.2017.

[3] Litman, T., Autonomous Vehicle Implementation Predictions. *Victoria Transport Policy Institute*, 2017.

[4] Neumann, A., *A paratransit-inspired evolutionary process for public transit network design*. Ph.D. thesis, TU Berlin, Berlin, 2014.

[5] Bischoff, J., M. Maciejewski, and K. Nagel, *City-wide Shared Taxis: A Simulation Study in Berlin*. submitted to ITSC 2017, 2017, available `http://www.vsp.tu-berlin.de/publications/vspwp/` as WP 17-11.

[6] Alonso-Mora, J., S. Samaranayake, A. Wallar, E. Frazzoli, and D. Rus, On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. *Proceedings of the National Academy of Sciences*, Vol. 114, No. 3, 2017, pp. 462–467.

[7] Xiao, J., H. Aydt, M. Lees, and A. Knoll, A partition based match making algorithm for taxi sharing. In *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, 2013, pp. 2393–2400.

[8] Hussain, I., L. Knapen, S. Galland, A.-U.-H. Yasar, T. Bellemans, D. Janssens, and G. Wets, Agent-based Simulation Model for Long-term Carpooling: Effect of Activity Planning Constraints. *Procedia Computer Science*, Vol. 52, 2015, pp. 412–419.

[9] Wagner, S., T. Brandt, and D. Neumann, Data Analytics in Free-Floating Carsharing: Evidence from the City of Berlin. In *2015 48th Hawaii International Conference on System Sciences*, 2015, pp. 897–907.

[10] Bischoff, J. and M. Maciejewski, Autonomous taxicabs in Berlin – a spatiotemporal analysis of service performance. *Transportation Research Procedia*, Vol. 19, 2016, pp. 176–186.

[11] *Inside Uber's quest to get more people in fewer cars*. Business Insider, 2016, `http://read.bi/1th2oFW`, accessed 17.07.2017.

[12] Uber, *It's a Beautiful (Pool) Day in the Neighborhood*, 2015, `https://newsroom.uber.com/us-california/its-a-beautiful-pool-day-in-the-neighborhood/`, accessed 07.07.2017.

[13] Davidson, A., J. Peters, and C. Brakewood, Interactive Travel Modes: Uber, Transit, and Mobility in New York City. In *TRB 96th Annual Meeting Compendium of Papers*, Transportation Research Board, Washington, DC, 2017.

[14] Shared-Use Mobility Center, *What Killed Kutsuplus? 3 Takeaways for Cities Pursuing Mobility-On-Demand*, 2016, `http://sharedusemobilitycenter.org/news/killed-kutsuplus-3-takeaways-cities-pursing-mobility-demand/`, accessed 17.07.2017.

[15] Horni, A., K. Nagel, and K. W. Axhausen (eds.) *The Multi-Agent Transport Simulation MATSim*. Ubiquity, London, 2016.

[16] Nagel, K. and G. Flötteröd, Agent-based traffic assignment: Going from trips to behavioural travelers. In *Travel Behaviour Research in an Evolving World – Selected papers from the 12th international conference on travel behaviour research* (R. Pendyala and C. Bhat, eds.), International Association for Travel Behaviour Research, 2012, pp. 261–294.

[17] Raney, B. and K. Nagel, An improved framework for large-scale multi-agent simulations of travel behaviour. In *Towards better performing European Transportation Systems* (P. Rietveld, B. Jourquin, and K. Westin, eds.), Routledge, London, 2006, pp. 305–347.

[18] Maciejewski, M., J. Bischoff, S. Hörl, and K. Nagel, Towards a Testbed for Dynamic Vehicle Routing Algorithms. In *Highlights of Practical Applications of Cyber-Physical Multi-Agent Systems: International Workshops of PAAMS 2017, Porto, Portugal, June 21-23, 2017, Proceedings* (J. Bajo, Z. Vale, K. Hallenborg, A. P. Rocha, P. Mathieu, P. Pawlewski, E. Del Val, P. Novais, F. Lopes, N. D. Duque Méndez, V. Julián, and J. Holmgren, eds.), Springer International Publishing, 2017, pp. 69–79.

[19] Maciejewski, M., Dynamic Transport Services. In (15), chap. 23.

[20] New York City Taxi and Limousine Commission, *2016 TLC Factbook*. New York City Taxi and Limousine Commission, 2016.

[21] Bischoff, J., M. Maciejewski, and A. Sohr, Analysis of Berlin's taxi services by exploring GPS traces. In *Models and Technologies for Intelligent Transportation Systems (MT-ITS), 2015 International Conference on*, 2015, pp. 209–215.