An Optimization-based Strategy for Shared Autonomous Vehicle Fleet Repositioning

Felipe de Souza[1], Krishna Murthy Gurumurthy[2], Joshua Auld[1] and Kara M. Kockelman[2]

[1]*Argonne National Laboratory, 9700 Cass Avenue, Lemont, IL, U.S.A.*

[2]*Department of Civil, Architectural and Environmental Engineering, University of Texas at Austin, Austin, TX, U.S.A.*

Keywords: Shared Autonomous Vehicles, Repositioning, Agent-based Simulation, POLARIS, Bloomington.

Abstract: With the emergence of autonomous technology, shared autonomous vehicles (SAVs) will potentially be the prevalent transportation mode for urban mobility. On one hand, relying on SAV fleets can provide several operational benefits. On the other hand, SAVs can increase travel distance and add congestion due to unoccupied trips such as pickup and repositioning trips. One important aspect for a SAV fleet's success is to serve the incoming requests at reasonably low waiting time. This is achieved by an adequate fleet size that is spatially distributed thoughtfully so that incoming requests can be served by a nearby vehicle. Unfortunately, it is challenging to keep a satisfactory spatial distribution of vehicles due to imbalances in the origin and destination patterns of incoming requests. This paper focuses on the impact of SAV relocation on traveler wait times using a novel optimization-based algorithm for repositioning. POLARIS, an agent-based tool, is used for a case study of Bloomington, Illinois to quantify the benefits of allowing SAV repositioning. On average, the wait times were around 20% lower with repositioning for all adequate fleet sizes. SAVs were available more uniformly across the region's zones, and proportional to trip-making at different times of day. In addition, enabling repositioning led to a higher share of demands being served. These benefits, however, are achieved at the expense of 6% added vehicles miles traveled.

1 INTRODUCTION

With the emergence of autonomous vehicles, travelers may relinquish their own private vehicles and rely on a fleet of shared autonomous vehicles (SAV's) (Spieser et al., 2014; Fagnant and Kockelman, 2014; Fagnant and Kockelman, 2015; Bischoff and Maciejewski, 2016; Stoiber et al., 2019) operating similarly to current Transportation Network Companies such as Uber and Lyft. On one hand, this shift can bring advantages such as the reduction of number of vehicles and lower the need for parking and garage spaces (Fagnant and Kockelman, 2015). On the other hand, the operation of SAV fleets can lead to significant unoccupied travel which can potentially worsen congestion, despite the increased capacity expected by autonomous vehicles (e.g., (Shladover et al., 2012)). Therefore, the operational strategies of SAV fleets should be such that they minimize empty travel while serving travelers in a timely manner.

Empty trips occur or empty vehicle miles traveled (eVMT) in the current human-driven TNC services, as well as with SAV fleets, in three different ways. First, the pickup trip for an assigned vehicle from its current location to the traveler location. Second, the trip performed at the beginning and end of a driver's shift from and to home, or a depot in the case of a SAV vehicle. Third, a repositioning trip from an area of low demand to an area with higher demand. In the case of human drivers, these trips are undertaken with the goal of reducing the time to serve a request and thereby increasing revenue at the expense of the additional empty trip's cost. In the case of autonomous vehicles, these trips occur to better serve future requests. There is some uncertainty on the share of empty travel in previous studies, and the extent to which dynamic ride-sharing can help. Recent literature has found that around 40% of the current TNC (Uber and Lyft) travel is empty (Henao and Marshall, 2019).These estimates are much lower for SAVs. Berlin study suggests the percent empty travel time as 17% (Bischoff and Maciejewski, 2016), whereas Austin scenarios, with and without pooling, averaged between 6 and 15% eVMT (Simoni et al., 2019; Gurumurthy et al., 2019).

Repositioning trips are important even though they add empty miles. The necessity of such trips

de Souza, F., Gurumurthy, K., Auld, J. and Kockelman, K.
An Optimization-based Strategy for Shared Autonomous Vehicle Fleet Repositioning.
DOI: 10.5220/0009421603700376
In *Proceedings of the 6th International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS 2020)*, pages 370-376
ISBN: 978-989-758-419-0

arise from an imbalance between the origin and destination locations of incoming requests. Areas that are common trip destinations (i.e, dropoff location) but not common origins accumulate vehicles while a dearth of vehicles is observed in areas that have many trip origins. The strategy proposed by (Alonso-Mora et al., 2017) found that 20% more trips can be served when repositioning is allowed. Another study proposed an assignment strategy that concurrently assigns vehicles to travelers while also dispatching vehicles to areas with high demand based on the expected future demand (Dandl et al., 2019). The share of repositioning miles ranged from 3 to 6% across all the simulation scenarios while the pickup miles remained around 12% of total miles.

In this study, the impact of SAV repositioning is studied at scale using the POLARIS agent-based framework. A computationally efficient repositioning strategy for SAV operation was implemented, and the operational results for the entire region of Bloomington, Illinois is discussed. The next section discusses the simulation framework, SAV modeling methodology and the repositioning algorithm. This is followed by results and discussions for the case study of Bloomington, Illinois, and finally ends with a conclusion.

2 POLARIS SIMULATION FRAMEWORK

POLARIS (Auld et al., 2016) is an agent-based framework for transportation systems that is designed to simulate large metropolitan areas. Elements of the simulated area and persons are modeled as individual agents that take decisions as the simulation evolves and is often refered to as a discrete event-based simulation. Specifically, POLARIS features an activity-based-model for the travel demand behavior, along with traffic flow and dynamic traffic assignment models. More recently, POLARIS also boasts a module for SAV simulation (Gurumurthy et al., 2020).

The travel behavior is captured by the ADAPTS activity-based model (Auld and Mohammadian, 2009; Auld and Mohammadian, 2012) which models various travel decisions ranging from within-day, mid-term and long-term choices. The mid-term and within-day travel behavior decisions include the process of individual activity episode planning and engagement. These decisions are constrained by long-term choices regarding home and workplace choice, and household vehicle choices, and, in turn, influence activity and travel planning.

Specific mode choices is modeled as an outcome of discrete choice models. POLARIS uses three separate mode-choice models depending on the activity purpose: home-based work/school, home-based other and non-home based, similar to traditional modeling. The nested-logit formulation used to model mode choice include nine modes: drive alone, TNC use, ride as passenger, walk, bike, bus with walk access, bus with drive access, rail with walk access, and rail with drive access. Among these, drive alone and TNC use are grouped under the auto nest, and the two rail modes are grouped under the rail nest. The model includes a variety of demographic variables, accessibility information, as well as level of service (LOS) variables. The demographic variables include individual demographics such as education, employment status, possession of driver's license, and household demographics such as household income, household size, and vehicle and bike ownership among others. Road-network density and activity density of the destination zone are used to capture the characteristics of land use and the transportation network. The TNC-specific Level of Service (LOS) variables used in the model include in-vehicle travel time and wait time (obtained from the simulation), and input fare. The TNC fare comprises of a fixed cost per trip, a distance-varying component, and a time-varying component. The model is developed and calibrated against the household travel survey data collected from the local region's metropolitan planning organization.

The realized travel times and delays along the simulation period is an outcome of traffic flow and dynamic traffic assignment models. The underlying traffic flow model is based on the link transmission models (Yperman, 2007) which in turn is based on Newell's kinematic wave model (Newell, 1993) with further adaptation to be able to track individual vehicles along their journey (de Souza et al., 2019). The dynamic traffic assignment algorithm (Auld et al., 2019a) assign routes to individual vehicles using a time-dependent A* shortest path router (Verbas et al., 2018) based on the prevailing traffic condition, as well as updated skim travel times. Traveler's routing behavior in response to delays is also captured by allowing re-routing.

The effects of real-time information and impacts of connected and automated vehicles - from both demand and supply sides - are also captured. This allows for exploratory studies on the impact of connected and automated vehicles in the overall transportation networks (Auld et al., 2019b), as well as the impact of shared mobility services in the presence of connected and automated vehicles (Gurumurthy et al., 2020).

3 SAV SIMULATION

The shared mobility simulation implemented in POLARIS (Gurumurthy et al., 2020) is extended here to evaluate the proposed repositioning strategy. SAVs are modeled to mimic operations that are currently observed in Transportation Network Companies' operation. The operator assigns requests to individual vehicles depending on the assignment strategy and monitors the spatial distribution of vehicles to determine repositioning decisions. SAVs execute the pickup, dropoff and repositioning tasks depending on the instruction received from the operator. SAVs are able to store requests that are being executed and those that need to be executed in the future. A detailed overview of the operator and vehicle operations is available at (Gurumurthy et al., 2020; de Souza et al., 2020).

3.1 Repositioning Strategy

The repositioning strategy aims to transfer vehicles from areas experiencing a high supply to areas experiencing a dearth of vehicles. This occurs due to an imbalance in the origin destination pattern of the incoming requests. Areas that are common trip destinations but not common origins tend to accumulate vehicles. Conversely, areas that are common trip origins experience vehicle shortage. Therefore, it is necessary to relocate idle vehicles into those areas in order to better serve future requests.

The repositioning strategy is based on zone-level variables. For each zone i, we define the supply s_i as the number of idle vehicles on zone i added of the number the non-idle vehicles whose the destination of last operation is at zone i (i.e., dropoff at zone i or repositioning to zone i), v_i as the number of idle vehicles at zone i. We also define f_i as the minimum supply at zone i. This minimum supply is an input to the proposed method. The minimum supply should roughly track the incoming demand for a given zone. In the following paragraphs we provide the basic guidelines to define f_i.

The goal of the strategy is to keep s_i higher than the minimum supply, f_i. The repositioning decision across the region is obtained through the following optimization problem:

$$\begin{aligned} \min_{x_{i,j}} \mathcal{J} = \sum_{\forall i}\sum_{\forall j} x_{i,j} t_{i,j} & \\ \sum_{\forall j} x_{j,i} - \sum_{\forall j} x_{i,j} + s_i \geq f_i & \qquad \forall i \\ \sum_{\forall j} x_{i,j} \leq v_i & \qquad \forall i \\ x_{i,j} \geq 0 & \qquad \forall (i,j), \end{aligned} \tag{1}$$

where $x_{i,j}$ is the number of vehicles that relocates from zone i to zone j and $t_{i,j}$ is the travel time from zone i to zone j or any zone-to-zone cost that is deemed appropriate. The optimization problem (1) is linear and therefore this problem can be efficiently solved with widely available solvers. Observe that variable $x_{i,j}$ must be integer as each unit is associated to a given vehicle repositioning from one zone to another. However, the constraints are unimodular and the costs are linear which guarantees that the solution of (1) yields discrete values as long as f_i is also integer. Unimodularity was also exploited in (Hyland and Mahmassani, 2018).

In addition, depending on the supply, s_i, the number of idle vehicles, v_i, a high f_i can turn the optimization problem infeasible. For example, if there is no idle vehicles (i.e., $v_i = 0 \quad \forall i$) and at least one zone in which $s_i < f_i$, the problem has no solution since there are no available vehicles to be relocated.

Therefore, there are two prerequisites for the minimum supply f_i when the optimization problem is instantiated: (i) it must be integer; (ii) it should be small enough so that the optimization problem is feasible. With respect to feasibility, we can evaluate whether a particular $\mathbf{f} = [f_1, ..., f_I]$ given $\mathbf{s} = [s_1, ..., s_I]$ and $\mathbf{v} = [v_1, ..., v_n]$ if:

$$\sum_{\forall i}(\max\{f_i - s_i, 0\}) \leq \sum_{\forall i}(\max\{\min\{s_i - f_i, v_i\}, 0\}), \tag{2}$$

that is, the term $\max\{f_i - s_i, 0\}$ yields the minimum number of vehicles that needs to be relocated to zone i while he right-hand-side accounts for the number of vehicles that zone i can supply to other zones.

We set f_i as follows. We assume a predicted demand d_i for zone i and we set f_i as:

$$f_i = \lfloor \alpha d_i \rfloor, \tag{3}$$

with the highest α in the interval $[0, 1]$ that satisfies the feasibility constraint (2) ($\lfloor x \rfloor$ is the integer part of x). Methods such as bisection can be used to find the highest α that satisfies the constraint. Here we perform a line search starting from $\alpha = 1$ and reducing with rate $\beta < 1$ as $\alpha := \beta\alpha$ until (2) is satisfied.

In summary, the following steps are performed:

1. Retrieve all s_i and v_i based on the current vehicles' statuses.
2. Obtain the predicted demand d_i (based on historical data or previous requests).
3. Obtain an α that satisfies (2).
4. Instantiate and solve the optimization problem (1).

5. Dispatch vehicles based on the solution $x_{i,j}$.

Here we assume the repositioning decisions are made at constant time steps (for example, every 5 minutes).

4 SIMULATION RESULTS

The repositioning strategy outlined above was tested for the Bloomington region in Illinois, USA. The network contains 185 zones, 7000 links, and 2500 nodes. The mode choice parameters are tuned to the current travel trends and yielded around 30,000 trip requests for the 24 hour simulation period. Three different fleet sizes of 650, 700, and 750 SAVs were tested with, and without, repositioning.

For all cases, the maximum waiting time is set to 10 minutes (i.e., the maximum pickup time for a SAV is 10 minutes as estimated before the start of the pickup trip. The realized travel time might be higher if traffic conditions changes). When repositioning is enabled, repositioning decisions are taken every 5 minutes, and d_i is obtained based on the number of requests on zone i in the previous 30 minutes of simulation. The GLPK solver (Makhorin, 2001) was used to solve optimization problem (1) at every step.

Table 1 presents the key metrics for each scenario. Without repositioning, the share of empty miles lies around 30% and it increased to around 33% in the cases in which repositioning was enabled. This increase in VMT allowed a lower share of the VMT in pickup trips since vehicles are repositioned to high demand areas. This allowed a average wait time 20% shorter when repositioning is available. Meanwhile, the share of served trips at peak time has increased for all fleet sizes.

The share of trips that were served and unserved over the 24 hr time period for all fleet sizes is depicted in Figure 1. Blue and orange lines correspond to the scenario without repositioning, and the green and red lines correspond to the scenario with repositioning. The fleet sizes are 650, 700, and 750 from the top to bottom. In all cases, enabling repositioning led to an increase in the share of served trips with the difference being larger for smaller fleets.

The repositioning method also lowers waiting time. Figure 2 depicts the distribution of waiting time for fleet sizes of 650 (top), 700 (middle), and 750 vehicles (bottom) with (in red) and without (in blue) repositioning. The vertical lines in red and blue highlight the average waiting time for each case, respectively.

The efficiency of an SAV fleet can be observed by assessing the daily operation profile. Figure 3 shows

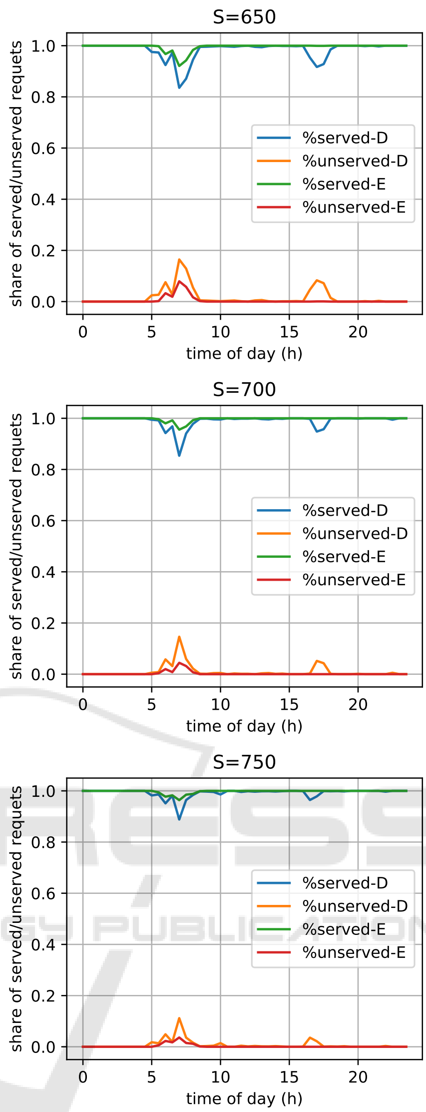

Figure 1: Share of Served and Unserved Trips for the Three Fleet Sizes with and without Repositioning Enabled.

the share of SAVs idle, or performing pickup, dropoff or repositioning for the two scenarios with and without repositioning. When repositioning is enabled, the entire fleet is utilized at peak times of day, if necessary. This is not true in the absence of repositioning when SAVs are idling at low demand areas.

5 CONCLUSIONS

An optimization-based method for SAV fleet repositioning is proposed. Vehicles in zones with excess supply are moved into areas where the supply is insufficient to serve the incoming demand. The method is formulated as a Linear Program with decision vari-

Table 1: Summary of the Results for the Three Different Fleet Sizes with and without Repositioning. Pickup VMT Is Labeled as pVMT, Repositioning Vmt as rVMT and Empty VMT as eVMT.

Fleet	VMT	% pVMT	% rVMT	% eVMT	% Served at Peak	Wait Time (min)
Without Repositioning						
650	205,475	30.3	0.0	30.3	83.5	4.52
700	206,403	29.6	0.0	29.6	85.3	4.32
750	202,513	28.3	0.0	28.3	88.8	4.07
With Repositioning						
650	221,843	24.3	9.2	33.5	92.0	3.49 (-23%)
700	223,002	23.3	10.0	33.2	95.5	3.37 (-22%)
750	220,431	22.5	10.5	33.1	96.3	3.24 (-20%)

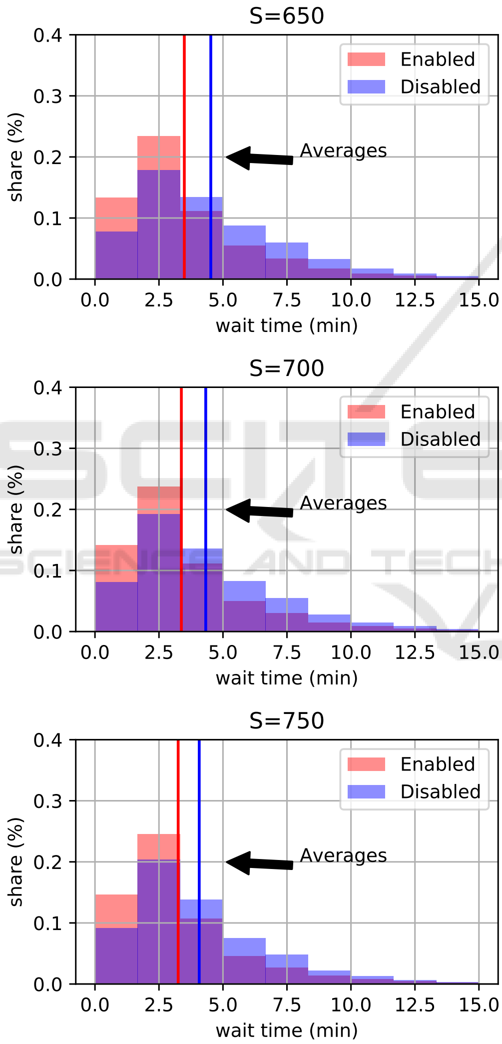

Figure 2: Histogram of Waiting Times for Different Fleet Sizes with (Red) and without (Blue) Repositioning.

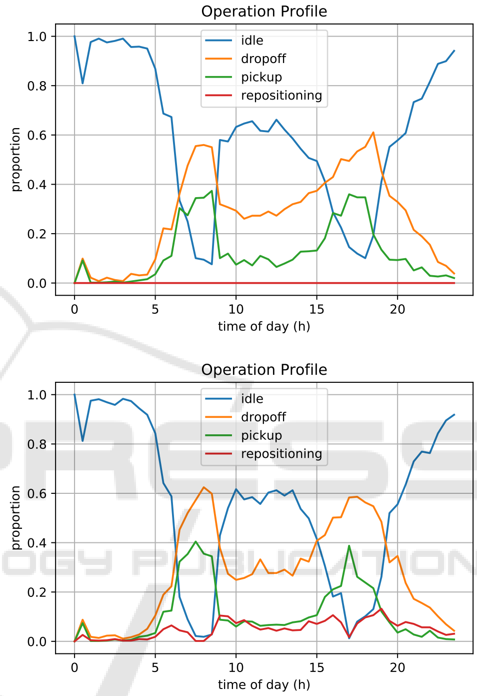

Figure 3: Fleet Operation Profile without (Left Graph) and with (Right Graph) Repositioning Enabled for $S = 650$.

ables based at the zone level. This brings two key advantages: (i) being a Linear Program, the method is computationally efficient; (ii) the computational complexity does not grow with the fleet size, since the number of vehicles is encoded as a variable as opposed to enumerating decisions specific for each vehicle in the network.

In experiments for a medium-sized network, enabling the repositioning strategy led to an increase in the share of served requests, as well as a reduction in waiting times of up to 20%. The benefits occurs at expense of higher empty distance traveled. The additional distance from repositioning partially balances with the added empty pickup distance, and, therefore, improves the fleet's service characteristics. In all, the results suggest the additional empty miles is significant smaller than the strategy introduced in (de Souza et al., 2020).

For future work, we plan to perform a thorough performance analysis using larger networks like that of the Chicago region. In addition, we want to investigate the influence of inaccuracies of the model inputs, especially with respect to the incoming demands, as well as the effect of different time steps for performing repositioning decisions.

ACKNOWLEDGEMENTS

This report and the work described were sponsored by the U.S. Department of Energy Vehicle Technologies Office under the Systems and Modeling for Accelerated Research in Transportation Mobility Laboratory Consortium, an initiative of the Energy Efficient Mobility Systems Program. David Anderson, a Department of Energy Office of Energy Efficiency and Renewable Energy manager, played an important role in establishing the project concept, advancing implementation, and providing ongoing guidance.

REFERENCES

Alonso-Mora, J., Samaranayake, S., Wallar, A., Frazzoli, E., and Rus, D. (2017). On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. *Proceedings of the National Academy of Sciences*, 114(3):462–467.

Auld, J., Hope, M., Ley, H., Sokolov, V., Xu, B., and Zhang, K. (2016). Polaris: Agent-based modeling framework development and implementation for integrated travel demand and network and operations simulations. *Transportation Research Part C: Emerging Technologies*, 64:101–116.

Auld, J. and Mohammadian, A. (2009). Framework for the development of the agent-based dynamic activity planning and travel scheduling (adapts) model. *Transportation Letters*, 1(3):245–255.

Auld, J. and Mohammadian, A. K. (2012). Activity planning processes in the agent-based dynamic activity planning and travel scheduling (adapts) model. *Transportation Research Part A: Policy and Practice*, 46(8):1386–1403.

Auld, J., Verbas, O., and Stinson, M. (2019a). Agent-based dynamic traffic assignment with information mixing. *Procedia Computer Science*, 151:864–869.

Auld, J. A., de Souza, F., Enam, A., Javanmardi, M., Stinson, M., Verbas, O., and Rousseau, A. (2019b). Exploring the mobility and energy implications of shared versus private autonomous vehicles. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pages 1691–1696. IEEE.

Bischoff, J. and Maciejewski, M. (2016). Simulation of city-wide replacement of private cars with autonomous taxis in berlin. *Procedia computer science*, 83:237–244.

Dandl, F., Hyland, M., Bogenberger, K., and Mahmassani, H. S. (2019). Evaluating the impact of spatio-temporal demand forecast aggregation on the operational performance of shared autonomous mobility fleets. *Transportation*, pages 1–22.

de Souza, F., Gurumurthy, K. M., Auld, J., and Kockelman, K. M. (2020). A repositioning method for shared autonomous vehicles operation. In *Presented at the 9th International Workshop on Agent-based Mobility, Traffic and Transportation Models, Methodologies and Applications (ABMTRANS), Warsaw, Poland*.

de Souza, F., Verbas, O., and Auld, J. (2019). Mesoscopic traffic flow model for agent-based simulation. *Procedia Computer Science*, 151:858–863.

Fagnant, D. J. and Kockelman, K. (2015). Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice*, 77:167–181.

Fagnant, D. J. and Kockelman, K. M. (2014). The travel and environmental implications of shared autonomous vehicles, using agent-based model scenarios. *Transportation Research Part C: Emerging Technologies*, 40:1–13.

Gurumurthy, K. M., de Souza, F., Enam, A., and Auld, J. (2020). Large-scale simulation of shared autonomous vehicles: Integrating the supply and demand perspectives. In *Presented at the 99th Annual Meeting of the Transportation Research Board (TRB), Washington DC*.

Gurumurthy, K. M., Kockelman, K. M., and Simoni, M. D. (2019). Benefits and costs of ride-sharing in shared automated vehicles across austin, texas: Opportunities for congestion pricing. *Transportation Research Record*, 2673(6):548–556.

Henao, A. and Marshall, W. E. (2019). The impact of ride hailing on parking (and vice versa). *Journal of Transport and Land Use*, 12(1).

Hyland, M. and Mahmassani, H. S. (2018). Dynamic autonomous vehicle fleet operations: Optimization-based strategies to assign avs to immediate traveler demand requests. *Transportation Research Part C: Emerging Technologies*, 92:278–297.

Makhorin, A. (2001). Gnu linear programming kit (glpk). *Department for Applied Informatics, Moscow Aviation Institute, Moscow, Russia*.

Newell, G. F. (1993). A simplified theory of kinematic waves in highway traffic, part i: General theory. *Transportation Research Part B: Methodological*, 27(4):281–287.

Shladover, S. E., Su, D., and Lu, X.-Y. (2012). Impacts of cooperative adaptive cruise control on freeway traffic flow. *Transportation Research Record*, 2324(1):63–70.

Simoni, M. D., Kockelman, K. M., Gurumurthy, K. M., and Bischoff, J. (2019). Congestion pricing in a world of self-driving vehicles: An analysis of different strategies in alternative future scenarios. *Transportation Research Part C: Emerging Technologies*, 98:167–185.

Spieser, K., Treleaven, K., Zhang, R., Frazzoli, E., Morton, D., and Pavone, M. (2014). Toward a systematic approach to the design and evaluation of automated mobility-on-demand systems: A case study in singapore. In *Road vehicle automation*, pages 229–245. Springer.

Stoiber, T., Schubert, I., Hoerler, R., and Burger, P. (2019). Will consumers prefer shared and pooled-use autonomous vehicles? a stated choice experiment with swiss households. *Transportation Research Part D: Transport and Environment*, 71:265–282.

Verbas, Ö., Auld, J., Ley, H., Weimer, R., and Driscoll, S. (2018). Time-dependent intermodal a* algorithm: Methodology and implementation on a large-scale network. *Transportation Research Record*, 2672(47):219–230.

Yperman, I. (2007). The link transmission model for dynamic network loading.