The Impact of Ridesharing in Mobility-on-Demand Systems: Simulation Case Study in Prague

Davide Fiedler¹, Michal Čertický¹, Javier Alonso-Mora², and Michal Čáp¹,²

Abstract—In densely populated-cities, the use of private cars for personal transportation is unsustainable, due to high parking and road capacity requirements. The mobility-on-demand systems have been proposed as an alternative to a private car. Such systems consist of a fleet of vehicles that the user of the system can hail for one-way point-to-point trips. These systems employ large-scale vehicle sharing, i.e., one vehicle can be used by several people during one day and consequently, the fleet size and the parking space requirements can be reduced, but, at the cost of a non-negligible increase in vehicles miles driven in the system. The miles driven in the system can be reduced by ridesharing, where several people traveling in a similar direction are matched and travel in one vehicle. We quantify the potential of ridesharing in a hypothetical mobility-on-demand system designed to serve all trips that are currently realized by private car in the city of Prague. Our results show that by employing a ridesharing strategy that guarantees travel time prolongation of no more than 10 minutes, the average occupancy of a vehicle will increase to 2.7 passengers. Consequently, the number of vehicle miles traveled will decrease to 35 % of the amount in the MoD system without ridesharing and to 60% of the amount in the present state.

I. INTRODUCTION

In densely-populated cities, private cars are an unsustainable mode of personal transportation. Parking capacity and road capacity in such environments are typically insufficient to accommodate the private car traffic. At the same time, roads and parking space are difficult to expand due to scarcity of free land. In result, many modern cities are plagued by congested roads, unavailability of parking, and high levels of air pollution.

Mobility-on-demand (MoD) systems represent an alternative to transportation by private vehicles. In an MoD system, when a passenger requests a transportation service, a particular vehicle is assigned to pick up the passenger and carry him or her to the desired destination, where the passenger is dropped off. Then, the vehicle is available again to serve other passengers. When an on-demand vehicle travels without a passenger, e.g., when it is moving from one drop-off position to the next pick up position, we say that it drives unallocated. Examples of on-demand systems include taxi service, transportation network companies such as Uber, or future systems of self-driving taxis being developed by companies such as Waymo, Uber or nuTonomy. The MoD systems employ massive vehicle sharing, and thus they can serve the existing transportation demand with smaller highly-utilized vehicle fleet and in turn, they promise to dramatically reduce the need for urban parking capacity. The performance of large-scale mobility-on-demand systems has been recently studied within newly developed mathematical models that indeed confirm the ability to reduce the number of vehicles in the system [1], [2]. However, they also show that the total amount of vehicle miles driven in the system is bound to increase due to unallocated vehicle trips [3]. For example, the case studies in Singapore [3] and Prague [4] indicate that unallocated traffic in the MoD system would add more than 30 % to total traffic in the system and subsequently would contribute to the formation of congestion [4].

A possible solution to the above problem is to implement large-scale ridesharing, where multiple passengers that travel in the same direction are matched and transported in one vehicle. Efficient ridesharing can increase vehicle occupancy and consequently reduce the required fleet size and the total amount of vehicle miles driven in the system.

The objective of this paper is to quantify the potential of ridesharing to reduce vehicular traffic in a large-scale MoD system. We use agent-based simulation to analyze hypothetical scenario of replacing all privately owned vehicles in the city of Prague by a) a single-occupancy MoD system and by b) an MoD system with ridesharing.

A. Related work

The impact of potential large-scale MoD system deployments have been studied extensively in recent years, both using simulation [1] and by means of formal mathematical modeling of such systems [3]. The impact on congestion arising from unallocated vehicle trips in MoD system has been numerically explored in [4]. One of the proposed mitigation is strategies is congestion-aware fleet routing [5].

Ridesharing was traditionally formalized in the framework of Vehicle Routing Problems (VRP) [6], typically as a specific variant of VRP with Pickup and Delivery [7] or a Dial a Ride Problem (DARP) [8]. Yet, the existing VRP methods focus on instances with tens of vehicles and requests [9] and as such, they are not applicable to management or analysis of large-scale fleets that often consist of thousands of vehicles and requests.

The potential for large-scale ridesharing within MoD systems was studied using the shareability network model in [10]. This analysis of taxi trips in Manhattan revealed that up to 80% of the trips could be pairwise shared such that the travel time is increased be no more than a couple of minutes. The analysis was later extended to other cities [11]. The assumption of maximum two passengers in a vehicle was later lifted in [12], where the authors proposed a

¹Department of Computer Science, Faculty of Electrical Engineering, CTU in Prague, Czech Republic david.fiedler@agents.fel.cvut.cz
²Department of Cognitive Robotics, 3ME, TU Delft, the Netherlands
technique for dynamic assignment of requests to vehicles in the fleet and applied the technique to analyze the potential of ridesharing within NYC taxi dataset. In contrast to the above work, that analyzed the scenario of replacing all taxis in Manhattan with a fleet of 3000 shared taxis, we analyze the scenario of replacing all private vehicles with a fleet of 50,000 shared MoD vehicles and specifically study how would large-scale ridesharing affect the utilization of road infrastructure.

The city of Prague was chosen for the case study partly because we have access to the demand model for the area and partly because it is a "regular" metropolis that shares many characteristics with other world cities and thus the results are more likely to generalize. This is in contrast with the previously considered urban areas such as Manhattan or Singapore that have an extremely high density of travel demand which may lead to overly-optimistic estimates of system performance.

B. Contribution

Our work extends the existing models of large-scale MoD systems with consideration for the potential of ridesharing. More specifically, using agent-based simulation, we quantify the potential of ridesharing to reduce traffic load and total vehicle miles traveled. We compare how would be the existing travel demand served a) in the current system with private vehicles, b) in the MoD system without ridesharing and c) in the MoD system with ridesharing. We run high-fidelity realistic-scale simulation of the three scenarios for the city of Prague. In particular, we use a) real road network topology, b) demand with realistic structure and realistic scale, and c) we simulate the whole lifetime of the vehicle, including empty trips.

Analyzing of the ridesharing in the context of on-demand systems is essential because it can answer the question of whether the MoD system can be, in fact, deployed without overloading the capacity of existing road infrastructure.

II. Background Material

When analyzing the traffic congestion on a road segment, the fundamental diagram of road traffic [13] is often used. This model relates traffic density to traffic flow. Traffic density is the number of vehicles per unit of distance on the road segment, while traffic flow is the number of vehicles passing a reference point per unit of time. When traffic density grows, traffic flow also increases until it reaches a tipping point after which the flow starts dropping due to congestion. The tipping point is known as the critical density [14]. The exact shape of the diagram is typically determined by fitting empirical data from real-world observations of vehicular traffic. Subsequently, we will use the traffic density as a measure of utilization of a particular road segment and the critical density value \( y_c = 0.08 \text{ vehicle m}^{-1} \) from [15]. The road segments with traffic density above critical density \( y_c \) will be referred to as congested road segments.

III. Methodology

In this section, we describe the methodology that allows us to quantify the potential of large-scale ridesharing in a hypothetical scenario of replacing all private cars in the city of Prague by an MoD system.

First, we use a travel demand model to generate a representative collection of trips that are currently realized by private cars in Prague during the morning peak. Then, we design an MoD system that is capable of serving the existing trips with required service quality. Further, we design a method that finds request-vehicle matching for ridesharing. Finally, we simulate the entire system in micro-simulation and gather data for subsequent analysis. We will now discuss each step in detail.

A. Input data

The set of trips that represent the transportation demand is generated by the multi-agent activity-based model of Prague and Central Bohemian Region [16]. In contrast to traditional four-step demand models [17], which use trips as the fundamental modeling unit, activity-based models employ so-called activities (e.g., work, shop, sleep) and their sequences to represent the transport-related behavior of the population. Travel demand then occurs due to the necessity of the agents to satisfy their needs through activities performed at different places at different times. These activities are arranged in time and space into sequential (daily) schedules. Trip origins, destinations and times are endogenous outcomes of activity scheduling. The activity-based approach considers individual trips in context and therefore allows representing realistic trip chains.

The model used in this work covers a typical work day in Prague and the surrounding Central Bohemian Region. The population of over 1.3 million is modeled by the same number of autonomous, self-interested agents, whose behavior is influenced by their sociodemographic attributes, current needs, and situational context. Individual decisions of the agents are implemented using machine learning methods (neural networks, decision trees, random forests, etc.) and trained using various real-world data sets, including census data, travel diaries, and other transportation-related surveys. Planned activity schedules are simulated and tuned, and, finally, their temporal, spatial, and structural properties are validated against additional historical real-world data (origin-destination matrices and surveys) using six-step validation framework VALFRAM [18], [19]. The model generates over three million trips by all modes of transport in one 24-hour scenario, out of which there are roughly one million trips by private vehicles. For the analysis, we select only the personal car trips during the morning peak, i.e., such that start between 06:30 and 08:00, which yields a collection of about 130 000 trips.

B. System model

We adopt a station-based design of the MoD system [3]. That is, we partition the city into \( n = 40 \) regions using k-means clustering over the demand data, and we assume
that there is a station at the center of each such region. The result of this process is shown in Figure 1. The number of regions was chosen such that the average travel time from a station to a passenger is below 3 minutes. Stations serve as temporary parking lots for idle vehicles and they also contain facilities such as refueling/charging and cleaning.

For simplicity, we assume that each station has sufficient stock of vehicles to cover the transportation demand originating in its region. Further, we assume that the stock of vehicles at each station is stabilized through a vehicle rebalancing process, that is, the stations with a surplus of idle vehicles continuously send empty vehicles to stations that have a shortage of idle vehicles. In particular, we use the rebalancing policy based on the solution of the optimal transport problem [20] used, e.g., in the Singapore MoD case study [3].

We experimentally determined that in order to be able to serve every request from the nearest station (without ridesharing), the MoD system requires the total of 51,951 vehicles.

The MoD system is modeled as follows. Let $V = 1, \ldots, m$ be the set of all vehicles in the system. The transportation requests arriving to the system are represented by a sequence $(t_1, o_1, d_1), (t_2, o_2, d_2), \ldots$, where $t_i, o_i,$ and $d_i$ are the announcement time, origin point, and destination point of request $i$ respectively. The $i$-th transportation request is revealed only at time $t_i$.

The state of a vehicle $v$ at a particular time point encodes its current position, the set of requests that are currently onboard the vehicle and the current plan of the vehicle. The plan of a vehicle is represented as a sequence of plan orders $\pi = z_1, z_2, \ldots$, where each order $z_i$ is either to pick up a request $r$, or to drop-off a request $r$. For a vehicle plan to be valid, for each onboard passenger, the plan must contain an order to drop-off this passenger and for every order to pickup request $r$, the plan must contain an order to drop-off the request $r$ later in the sequence.

The operational cost of vehicle $v$ when following plan $\pi$ is denoted $s_v(\pi)$. For simplicity, we define $s_v(\pi)$ to be equal to the distance driven by the vehicle when it follows plan $\pi$. The travel delay (or discomfort) of request $r$ when the request is served by vehicle $v$ following plan $\pi_v$ is denoted $q_r(\pi_v)$. We define the travel delay as the difference between the travel time in a shared vehicle and a travel time along the fastest route:

$$q_r(\pi_v) := (t^{\text{dropoff}}_r - t_r) - t^{\text{baseline}}_r,$$

where $t^{\text{dropoff}}_r$ is the time when the request was dropped off under plan $\pi_v$, and $t^{\text{baseline}}_r$ is the duration of the fastest route from pickup point of request $r$ to the drop-off point of request $r$.

We desire to minimize the total operational cost of the system, such that the discomfort of every passenger is bounded by $q_{\text{max}}$. That is, we will attempt to minimize $\sum_v s_v(\pi_v)$ subject to $q_r(\pi_{v(r)}) \leq q_{\text{max}} \ \forall r$, where $\pi_{v(r)}$ is the plan of the vehicle that serves the request $r$.

### C. Request-vehicle matching

To implement ridesharing in an MoD system, one has to determine how should be the arriving travel requests assigned to individual vehicles in the system such that the overall distance driven in the system is minimized. This problem can be mathematically modeled as a dynamic dial-a-ride problem (D-DARP), which is known to be NP-hard [21]. Despite recent algorithmic advancements [22], [12], [23], the scalability of exact methods remains limited to problem instances of moderate size. Therefore, in this work, we compute distance-minimizing request-vehicle ridesharing assignments using an Insertion Heuristic [24]. The request-vehicle matching is implemented as follows:

When a new request $r = (t, o, d)$ arrives, the request-vehicle matching algorithm attempts to incorporate the request into the current plan of every vehicle. For a particular vehicle $v$, we try all possible points to insert pickup and drop-off order into existing plan, such that the relative ordering of the previous orders in the plan remains unchanged. Further, we measure the increase of operational cost for each plan generated by this process and select the plan (and the corresponding vehicle) that minimizes the increase in operational cost and at the same time satisfies the service discomfort constraints. The maximum discomfort threshold $q_{\text{max}}$ is chosen such that, if no other suitable ridesharing match is found, the request can always be served with delay smaller than $q_{\text{max}}$ using a vehicle dispatched from the nearest system station. For the pseudo-code of the request-vehicle matching method, see Algorithm 1.

In order to evaluate the operational cost and service discomfort associated with a particular plan, one needs to estimate the travel time of the vehicle from one order in the plan (e.g., from pickup position of one request) to the subsequent order in the plan (e.g., the drop-off position of some other request). Notice that the number of evaluation of travel time estimate can be up to the order of $O(m \cdot d^3)$. The fleet size in our case study will be $m \approx 50,000$ and the average length of the plan will be $l \approx 4$ resulting hundreds of thousands of evaluations of the travel time estimate within single request assignment procedure. In order to maintain computational tractability, instead of planing the
Algorithm 1: Finding a vehicle and a plan to serve request \( r \) using insertion heuristic.

1. **On arrival of new request \( r \)**
   2. for \( v \in V \)
      3. let \( \pi_v \) be the current plan of vehicle \( v \);
      4. let \( l \) be the length of plan \( p_v \);
      5. for \( i = 1, \ldots, l + 1 \)
         6. for \( j = i + 1, \ldots, l + 2 \)
            7. \( \pi_i' \leftarrow \text{insert pick-up}(r) \) order before position \( i \) in plan \( \pi_v \);
            8. \( \pi^{i,j}_v \leftarrow \text{insert drop-off}(r) \) order before position \( j \) in plan \( \pi' \);
            9. \( v^*, i^*, j^* \leftarrow \arg\min_{v,i,j} s_v(\pi^{i,j}_v) - s_v(\pi_v) \) subj. to \( \pi^{i,j}_v \) is valid and \( q_r(\pi^{i,j}_v) \leq q_{\max} \forall \text{req.}\ r \text{ served by plan } \pi^{i,j}_v ; \)
            10. request \( r \) is assigned to vehicle \( v^* \);
      11. vehicle \( v^* \) follows plan \( p_{v^*}^{i^*, j^*} ; \)

fastest route on the road network, we approximate the travel time between two locations on the map with a linear model based on the Euclidean distance between the two points. The model is calibrated to minimize the difference between the estimate and the duration of the fastest route in the road network. Since the road network travel time may be longer than the estimate, a plan that satisfies discomfort constraints computed using estimate may violate the constraint when evaluated using road network travel time. Nevertheless, in our experiments, we observed that the average road network travel delay is significantly lower than the maximum delay constraint.

D. Simulation

The scenarios were simulated in multi-agent transportation simulation framework AgentPolis\(^1\). AgentPolis is a large-scale multi-agent discrete-event simulation written in Java. The simulation environment consists of a) road network that is in turn composed of nodes (crossroads) connected by edges (road segments), b) stations in which on-demand vehicles park, c) on-demand vehicle agents, and d) passenger agents. In Figure 2, we show the city of Prague simulated in AgentPolis.

The topology of the road network is obtained from OpenStreetMap (OSM), resulting in a road network consisting of 158,674 edges and 63,995 nodes. The speed limit for each road segment was also taken from OpenStreetMap data and missing entries were generated according to following rules based on the local legislation: highway: 130 km/h, living street: 20 km/h, otherwise: 50 km/h.

During simulation initialization, we create the vehicles representing the on-demand fleet and assign them to stations. For each travel request, at the request announcement time, we create a passenger agent. The passenger is then picked up by the assigned on-demand vehicle, driven to the desired location, dropped off and finally released from the simulation. The vehicle to serve the passenger is selected using the request-vehicle matching procedure described in Section III-C. Note that each passenger can be either matched to one of the empty vehicles parked in a station or to a non-idle vehicle that is already on its way to serve previously assigned requests. After the passenger has been dropped off, the vehicle continues executing its plan. In the case the plan of the vehicle becomes empty, the vehicle returns to the nearest system station.

In this paper, we analyze the system during the morning peak, i.e., during the period between 7:00 and 8:00. To avoid the “cold start” artifacts, the simulation begins at 6:30, but for subsequent analysis, we only use the data generated between 7:00 and 8:00.

IV. RESULTS

In this section, we report on the results of our simulation analysis. We run the following three scenarios:

1) **Present situation scenario**: Each travel request is realized by a private vehicle.

2) **MoD scenario**: Travel demands is served by the MoD system.

3) **MoD with ridesharing scenario**: Travel demand is served by the MoD system with ridesharing. Request-vehicle matching uses maximum travel delay \( q_{\max} = 10\ min \).

Figure 3 shows the distribution of the traffic density over the road network for the three compared scenarios. We can see that when the transportation demand is served by an MoD system without ridesharing, the traffic density at most parts of the road network increases compared to the present situation. When the MoD system employs ridesharing, however, the traffic density at most road segments decreases relative to both the present situation and to the MoD system without ridesharing.

Figure 4 shows the histogram of travel densities in the three scenarios. The first row shows the traffic density histograms for all edges with non-zero density (used at least

\(^1\)https://github.com/aicenter/agentpolis

Video: https://sum.fel.cvut.cz/itsc2018/

Fig. 2. AgentPolis visualization of the traffic in Prague at 07:00.
Fig. 3. Comparison of traffic densities over the road network in the three compared scenarios. The darker colors indicate higher traffic density, black color is reserved for the roads that are congested, i.e., the traffic density exceeds the critical density.

<table>
<thead>
<tr>
<th></th>
<th>Present</th>
<th>MoD</th>
<th>MoD with ridesharing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total veh. dist. traveled (km)</td>
<td>940 645</td>
<td>1 586 495</td>
<td>560 875</td>
</tr>
<tr>
<td>Avg. density (veh/km)</td>
<td>0.0080</td>
<td>0.0101</td>
<td>0.0052</td>
</tr>
<tr>
<td>Congested segments</td>
<td>14</td>
<td>55</td>
<td>4</td>
</tr>
<tr>
<td>Heavily loaded segments</td>
<td>208</td>
<td>551</td>
<td>35</td>
</tr>
</tbody>
</table>

**TABLE I**

**COMPARISON BETWEEN THE THREE CONSIDERED SCENARIOS.**

**Congested segments** are segments on which traffic density is above critical density, **heavily loaded segments** are segments with density above 50% of the critical density.

once in the analyzed time window) while the second row shows histograms that are “zoomed-in” to the values around the critical density. Again, we can see that MoD system with ridesharing is able to serve the current transportation demand using less road capacity than the private vehicles use today and than the on-demand vehicles would use if they do not employ ridesharing.

Table I summarizes the performance of the three evaluated systems. It is remarkable that ridesharing can reduce the total amount of vehicle miles traveled in an MoD system (and thus also its energy consumption) almost three-fold.

To understand the effectiveness of ridesharing, the vehicle occupancy is one of the most important indicators. The occupancy of each non-idle vehicle (i.e., a vehicle that is not in a station) was measured during the simulation in one-minute intervals. Figure 5 shows a histogram of vehicle occupancies measured in period 7:00 - 8:00. The MoD system with ridesharing has an average occupancy of 2.74 passengers per vehicle. This represents a significant improvement over the MoD system without ridesharing that has the average occupancy of 0.7 passengers per vehicle.

In order to get better insights into the relationship between the travel discomfort and vehicle occupancy, we simulated MoD system with ridesharing and travel delay constraints $q_{max} = 7, 10, 12$ and 15 minutes. Figure 6 shows the relation between the maximum delay and the average vehicle occupation. As expected, the vehicle occupancy can be increased, and in turn, operation cost reduced if we sacrifice the comfort of passengers.

**V. CONCLUSION**

In cities, the private vehicle is becoming an unsustainable mode of personal transportation due to its significant parking and road capacity requirements. Mobility-on-Demand (MoD) systems represent an alternative to transportation by private vehicles. These systems employ massive vehicle sharing, and thus they can serve existing transportation demand with a smaller highly-utilized fleet and in turn drastically reduce the need for parking capacity. However, studies show that unallocated trips in such systems would significantly increase vehicular traffic compared to the present state.

A proposed solution to this problem is to employ large-scale ridesharing. The potential of this solution strategy however remained unclear.

In this paper, we quantified the potential of ridesharing to reduce vehicular traffic in an MoD system. We analyzed hypothetical scenario of MoD deployment in the city of Prague and compared performance of three scenarios: a) the current state, where the demand is served by private vehicles, b) an MoD system without ridesharing and c) an MoD system with ridesharing.

The simulation results demonstrate that the MoD systems with ridesharing can not only compensate for the traffic generated by the unallocated trips within on-demand systems, but that ridesharing can significantly improve the road network utilization even with respect to the present state. In particular, by employing ridesharing, the average total traveled distance was reduced to 35% of the distance traveled in MoD system without ridesharing and to 60% of the distance traveled by private vehicles in the present state.

In future, we plan to study the applicability of more sophisticated request-vehicle matching methods to such large-scale scenarios and conduct test with more complex congestion models.

**Acknowledgments:** This research was supported by the Czech Science Foundation (grant No. 18-23623S), AMS Institute, and OP VVV MEYS funded project CZ.02.1.01/0.0/0.0/16_019/0000765 "Research Center for Informatics". Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum, provided under the program "Projects of Large Infrastructure for Research, Development, and Innovations" (LM2010005), is greatly appreciated.
Fig. 4. Histogram of traffic densities in the three different scenarios – from the left: private vehicles, MoD system, and MoD system with ridesharing. In the second row, only the bins with traffic density above 50% of critical density (0.04 vehicle m⁻¹) are shown to improve readability. Edges with zero traffic density were discarded from all histograms.

Fig. 5. Vehicle occupancy histogram

Fig. 6. The relation between the max. travel delay in minutes (horizontal axis) and the avg. vehicle occupancy (vertical axis)

REFERENCES