ANALYZING THE DYNAMIC RIDE-SHARING POTENTIAL FOR SHARED
AUTONOMOUS VEHICLE FLEETS USING CELLPHONE DATA FROM ORLANDO,
FLORIDA

Krishna Murthy Gurumurthy
Department of Civil, Architectural and Environmental Engineering
The University of Texas at Austin
gkmurthy10@utexas.edu

Kara M. Kockelman, Ph.D., P.E.
(Corresponding Author)
Professor and E.P. Schoch Professor in Engineering
Department of Civil, Architectural and Environmental Engineering
The University of Texas at Austin – 6.9 E. Cockrell Jr. Hall
Austin, TX 78712-1076
kkockelm@mail.utexas.edu

Under review for presentation at the 97th Annual Meeting of the Transportation Research Board in January, 2018 and for publication in Transportation Research Record.

ABSTRACT
Transportation network companies (TNCs) are regularly demonstrating the economic and operational viability of dynamic ride-sharing (DRS) to any destination within a city, thanks to real-time information from smartphones. In the foreseeable future, fleets of shared automated vehicles (SAVs) may largely eliminate the need for human drivers, while lowering per-mile operating costs and increasing the convenience of travel. This may dramatically reduce private vehicle ownership and deliver extensive use of SAVs. This study anticipates DRS matches across different travelers by using AirSage’s cellphone-based trip tables across 1,267 zones over 30 days. Assuming that the travel patterns do not change significantly in the future, the results suggest significant opportunities for DRS-enabled SAVs. Nearly 70% of the single-person trips could be shared with other persons traveling solo and with less than 5 minutes added travel time (to arrive at their destinations), and this value climbs to 90% for 15 to 30 minutes of added wait or travel time.

INTRODUCTION
Traffic safety and congestion are key transportation issues for many regions around the world. Driver error remains the dominant reason for vehicle crashes (NHTSA, 2015) and rising vehicle-miles traveled (VMT) is not worsening traffic congestion (FHWA, 2017). The introduction of autonomous vehicles (AVs) for personal use may dramatically reduce vehicle collisions by eliminating driver error. AVs will also improve mobility options for many travelers, especially those without driver’s licenses.

Several transportation network companies (TNCs) offer a dynamic ride-sharing (DRS) option, like Uber Pool and Lyft Line. Ride-sharing is not a new concept (Chan and Shaheen, 2012), with carpooling often feasible for those with common origins and destinations, and stable, similar
departure times on both ends of a round-trip (e.g., for many school trips within a neighborhood and for certain work trips). In practice, only casual carpooling or ‘slugging’ tends to serve real-time demands of flexible departure times (Ma and Wolfson, 2013; Dai, 2016), and is limited to very special corridors (where high toll and time savings induce many drivers to open their doors to different, unknown passengers every day).

Smartphone technology is fundamental to more widespread use of DRS, since it enables real-time access to traveler (and vehicle) locations (Amey et al., 2014). Exploiting this feature, TNCs have designed user-friendly car-sharing platforms that interface passengers and drivers, at any time of day and in any region the TNCs serve. By selecting the DRS option, travelers costs (but not travel times) are lowered, thanks to TNCs working to match two or more travelers with overlapping real-time routes. Such matches add some travel time, but deliver significant trip-cost savings and often good conversations among those sharing the ride, who had been strangers (alongside a TNC driver also on board).

AVs will be expensive, at least initially, and not be available for personal ownership for many years (Bansal and Kockelman, 2017). Fleet operators may profitably invest in a fleet of AVs, and manage them as TNCs currently manage their (driver-supplied) fleets, but with lower labor costs and complete control of plans and routes. Safer technologies should eventually bring down insurance costs, making shared AVs, or SAVs, more economically viable. In terms of congestion, SAVs offering DRS can increase average vehicle occupancy (AVO) and reduce regional VMT (Fagnant and Kockelman, 2016; Rodier et al., 2016). It is useful to quantify the level of opportunity for such services, across a range of settings.

This paper studies the DRS potential of for trip-making across the City of Orlando, Florida, as serviced by a fleet of SAVs. It relies of trip tables derived from cellphone data, as provided by AirSage across a period of 30 consecutive days, to provide a sense of day-to-day trip-making variations. The remaining paper summarizes related work, describes the AirSage data set, and then explains the methodology used to match distinct vehicle trips or traveling parties. All simulation results are presented, along with various conclusions.

RELATED LITERATURE

Over the past 10 years, several contributions have been made to optimize and/or implement DRS, with various researchers suggesting that DRS is a key method for reducing future roadway congestion (Levofsky and Greenberg, 2001; Berbeglia et al., 2010; Ma et al., 2013; Farhan and Chen, 2017; Levin et al., 2017). More recently, DRS has been successfully demonstrated using agent-based models (see, e.g., Fagnant and Kockelman, 2016; Bischoff et al., 2016; Loeb et al., 2017; and Hörl, 2017). Horni et al. (2016) used MATsim and a synthetically generated dataset of people and journeys to simulate dynamic traffic conditions.

When it comes to actual trip-making, mode choices, and traffic patterns, DRS has been investigated for cities like Atlanta, Georgia, Taipei, Taiwan, and New York City. DRS applications include the entire U.S. state of New Jersey and the nation and Singapore, using travel demand model trip-making predictions, publically available taxi datasets, and/or synthetically generated itineraries. Investigations demonstrate system feasibility and/or assess the computational
efficiency of different methods for assigning vehicles and/or matching travelers in shared rides. (See Agatz et al., 2011; Santi et al., 2014; Alonso-Moro et al., 2016; Brownell and Kornhauser, 2014; Bhat, 2016; and Tao, 2007.)

Agatz et al. (2011) developed a sophisticated algorithm to match riders to their drivers and conducted a simulation using person-trip data obtained from Atlanta’s travel demand model. Their results suggest that DRS works well not only in high-density, high-use settings, but in sprawling suburbs and at low rates of utilization. However, they focused on driver (and thus TNC vehicle) availability, which can hamper sharing and dilute DRS opportunities. Brownell and Kornhauser (2014) focused on SAV system performance for the state of New Jersey. Employing a gridded-network for the entire state, along with synthetic trip-making data, valuable precision, accuracy, and applicability may have been lost in assessing optimal fleet requirements.

Santi et al. (2014) and Alonso-Moro et al. (2016) overcame both these issues by using publicly available taxi datasets for New York City and real networks (via OpenStreetMaps, an open-source platform for map data). Alonso-Moro et al. observed that 98% of the City’s taxi trips could be served with just 2,000 vehicles and low waiting times (averaging just 2.8 minutes). Bhat’s (2016) confirmed those New York City taxi results, and added a vehicle repositioning algorithm. Tao (2007) also used a taxi data set, but for the city of Taipei. He developed a heuristic DRS algorithm using real-time taxi movements (not just trip calls by travelers) to test its efficiency in a realistic network setting. Tao (2007) achieved 60% ride matches and concluded that a higher matching rate could be obtained across larger networks with greater density of trip-making.

Of course, taxis do not represent all person-trips in any region. Such trips tend to be shorter than household-vehicle trips (due to their cost), more often for business reasons or those without parking access (again due to their cost), and for visitors (due to their unfamiliarity with the region). DRS investigations of more representative trip-making are desired. By using a population-weighted cellphone dataset, as done here, one overcomes the drawbacks of faked or taxi-based trip patterns. However, certain details are lost (such as trip-to-trip connections throughout the day), in order to protect travelers’ privacy, over space and time. Thus, cell-phone-based trips or other forms of extensive diary data tend to be aggregated by traffic analysis zones (TAZs) or neighborhoods, to obscure home and work addresses. To keep data size manageable (for dataset sharing), trips are often aggregated into hourly or multi-hour time-of-day bins as well. More detailed trip ends and trip schedules can be simulated/faked and disaggregated, while preserving the population’s basic trip patterns. This process ensures that matches are less obvious (with trips coming from all over a zone and hour, rather than from its centroid or mid-point, for example), and so was used here. But it comes at the expense of some accuracy and precision (versus the reality of actual trip locations and times, which are rarely available to anyone, for any large population).

CELLPHONE DATASET

The cellphone-based dataset employed here was generated by AirSage for the month of April 2014 and for travel across the City of Orlando, Florida. AirSage uses the regular location pings of cell phones that are turned on and carried by customers of its partner companies (like Verizon and Sprint). Cellphone trips observed were aggregated based on seven factors: each trip’s inferred
origin and destination TAZs, the hour and day in which most of the trip was made (e.g., 0100-0200 on April 4 or 1600-1700 on April 20), inferred trip purpose, and cell-phone subscriber class. All trips (and basic demographics) inferred from phone pings (of the carriers’ cell towers) were then expanded to reflect all trip-making in the City (including travel by persons who do not own cell phones or carry theirs with them, turned on).

The Orlando region’s metropolitan planning agency models travel across 1,267 TAZs (with 1,261 of them representing just the City of Orlando and the remaining 6 representing external counties). External-zone trips can be very long, with ambiguity in their true destination or origin, so all external trips were removed from the dataset before seeking matches. Traveler type also is not relevant, so it is not used here, in making matches (though one can imagine a future where some women may prefer to share rides with other women, and/or some people may prefer matches made within their age group, for example).

MetroPlan Orlando, the region’s metropolitan planning organization (MPO), provided a detailed network, with nearly 24,000 nodes and around 61,000 links. Shortest-path travel times (for each broad time of day, to reflect congested and congested conditions at peak and off-peak times of day) between each TAZ were used while disaggregating the trips, as discussed in the next section.

**METHODOLOGY**

**Data Disaggregation**

Since AirSage provided an anonymized, spatially and temporally aggregate dataset (with trips classified into hourly bins and their origins and destinations by TAZs), smaller time steps and more detailed locations were needed for a DRS application of intra-regional trips. A time-step of one minute was used here, to facilitate computation while preserving dataset integrity. Thus, the original 30-day 24-hour dataset was disaggregated into 30 different sets of 1440 (60 min/hr x 24 hr) one-minute trip-request files.

The departure times of these trips were not always in the hourly bin that Airsage indicated for each trip, because trips (within this region) can begin 15 minutes earlier (or can end 15 minutes later). This is because only the majority of the trip’s duration had to have occurred in the hour bin to which the trip was assigned by Airsage. (Few trips in the Orlando region are over 30 minutes in duration, so few could have started earlier or ended later than this.) Once a start time was assigned, the shortest-path travel times for that time of day, as obtained via Caliper Corporation’s TransCAD software, a travel-demand modeling tool, were used to sample individual trip travel times from a normal distribution, whose mean equaled this shortest-path travel time. The results were smooth, minute-by-minute trip-request files for each of the 30 days, and natural looking departure and arrival time patterns throughout each of the 30 days.

**Day to Day Variability in Travel Patterns**

The cumulative trip distribution for each of the 30 days was obtained by time of day, as shown in Figure 2. Variability, and consequently correlation, between each day was assessed using R software’s statistical tool. Table 1 shows correlation coefficients for trip counts across all origin-destination pairs and across all 30 days of the month, with shading to highlight correlation patterns.
Table 1 indicates that high correlation exists for trip patterns on Saturdays and Sundays, and for those made on weekdays, as one would expect (since weekdays have high shares of work and school trips, starting early in the day, while weekends have more flexible departure times and more recreational trip-making). Given these similarities, the following results are presented for a single weekday and a single weekend day. Results are very similar for other days of the 30-day dataset.

![Trip Distribution Chart](image)

**FIGURE 1** Difference in trip distributions by time of day between weekdays and weekends.

**Trip Matching Results**

An analysis of these trip patterns suggests how many trips can be matched with other trips, enabling ride-sharing, under different trip-delay and re-routing assumptions. MATLAB code was developed to identify trips whose rides (in an SAV, for example) can be shared. An assumption of 4-person maximum vehicle occupancy was made, along with various travel delay thresholds, before running the code, for various maximum-delay scenarios (ranging from 5 minutes of extra travel time, to a maximum of 30 minutes).

Figure 2 illustrates how travel times under ride-sharing is calculated for this exploratory analysis, with ride-sharing en route (not just for those sharing an origin zone and an destination zone, while having similar departure times). Including the entire dataset of trips would mean that trips that are already shared/performed together, like family members travelling together for dinner, inflate the trip-sharing percentages. Florida DOT (2013) estimates that over 50% of all automobile trips in that state are driven alone. Thus, only a fraction the person-trips in the Airsage dataset were used here, to perform matching (of solo travelers with one another, rather matching those already in traveling parties).
### TABLE 1 Correlation between Trips Made on All Days for the Month of April

<table>
<thead>
<tr>
<th>Day</th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thu</th>
<th>Fri</th>
<th>Sat</th>
<th>Sun</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thu</td>
<td>0.9908</td>
<td>0.9954</td>
<td>0.9593</td>
<td>0.9598</td>
<td>0.9954</td>
<td>0.9650</td>
<td>0.9978</td>
</tr>
<tr>
<td>Fri</td>
<td>0.9979</td>
<td>0.9935</td>
<td>0.9593</td>
<td>0.9598</td>
<td>0.9954</td>
<td>0.9650</td>
<td>0.9978</td>
</tr>
<tr>
<td>Sat</td>
<td>0.9935</td>
<td>0.9593</td>
<td>0.9568</td>
<td>0.9593</td>
<td>0.9954</td>
<td>0.9650</td>
<td>0.9978</td>
</tr>
<tr>
<td>Sun</td>
<td>0.9935</td>
<td>0.9593</td>
<td>0.9568</td>
<td>0.9593</td>
<td>0.9954</td>
<td>0.9650</td>
<td>0.9978</td>
</tr>
</tbody>
</table>
FIGURE 2 Illustrations of Fleet-sharing with and without DRS.

RESULTS
As shown in Table 3, even after removing a large share of trips that reflect traveling parties (and thus focusing only on Orlando trips undertaken by a single person), nearly 70% of all such single-person trips can be shared with less than 5 minutes of added total travel (for each of the ride-sharing travelers, including any wait time added). This percentage reaches 95% matching or shared when travelers are willing to wait (or delay their destination arrivals, for example) up to 30 minutes. Of course, not all travelers need to be willing to wait that long; most of the matches are made with added delays of under 5 minutes.

TABLE 3 Percentage of Trips That Can Be Shared With and Without DRS for a 4-passenger SAV under Different Maximum-Delay Assumptions

<table>
<thead>
<tr>
<th>Maximum Added Travel Time (including wait time)</th>
<th>Percentage of Trips that Can be Shared</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 min</td>
<td>69.3%</td>
</tr>
<tr>
<td>10</td>
<td>84.3%</td>
</tr>
<tr>
<td>15</td>
<td>90.1%</td>
</tr>
<tr>
<td>20</td>
<td>93.1%</td>
</tr>
<tr>
<td>25</td>
<td>94.8%</td>
</tr>
<tr>
<td>30</td>
<td>95.5%</td>
</tr>
</tbody>
</table>

CONCLUSIONS
This study anticipates the fraction of single-person trips that appear easily matched with one another, making them excellent candidates for dynamic ride-sharing across the City of Orlando. Several studies have simulated the operations of SAV fleets but without the comprehensive nature of this cellphone-based dataset (e.g., taxi datasets do not reflect other modes of travel) and/or without other key data (e.g., actual travel times). With such data in hand, and a new setting for simulation (a Florida city and major destination for many vacationers), the results obtained here
may be relevant for many interested in encouraging SAV use and DRS, to keep travel costs, VMT, emissions, and congestion down, as self-driving vehicles start making travel easier.

The trip-matching algorithm employed here suggests that nearly 70% of all single-person trips occurring each day in Orlando appear matchable to other trips taking place (for those traveling solo), with less than 5 minutes of added total travel time (including any wait time). Any added willingness to wait (up to 10 minutes or 15 minutes, maximum, for example) brings this percentage up (to 84.3% and 90.1%, respectively), suggesting substantial opportunities for VMT reduction and shared-fleet activities in many (and probably all) cities around the U.S. and presumably around the world. Coming research will disaggregate the trips spatially and devise new fleet-management algorithms for a shared fleet of vehicles, to enable more accurate and detailed performance metrics (like wait times and fleet VMT) for DRS across Orlando.

ACKNOWLEDGEMENTS

We thank Mr. Vijay Sivaraman, from AirSage, for providing the robust cellphone dataset that enabled this analysis. We also thank the MetroPlan Orlando staff, namely Gary Huttman, Nick Lepp and Nikhila Rose, for describing current travel patterns and sharing their regional travel demand model and network. This research effort was financially supported by TxDOT project 0-6838, with super-computing resources from the Texas Advanced Computing Center (TACC). The writers also thank Scott Schauer-West and Amy Banker for their edits and administrative support.

REFERENCES


