THE INTERPLAY BETWEEN FLEET SIZE, LEVEL-OF-SERVICE AND EMPTY VEHICLE REPOSITIONING STRATEGIES IN LARGE-SCALE, SHARED-RIDE AUTONOMOUS TAXI MOBILITY-ON-DEMAND SCENARIOS

Shirley Zhu
Department of Operations Research and Financial Engineering
Princeton University
229 Sherrerd Hall
Princeton, NJ 08544
shirleyzhu94@gmail.com

Alain L. Kornhauser, Ph.D
Professor, Department of Operations Research and Financial Engineering
Princeton University
229 Sherrerd Hall
Princeton, NJ 08544
alaink@princeton.edu
Phone: 609-258-4657

Word Count: 3792 words + 4 figure(s) + 4 table(s) = 5792 words
ABSTRACT

The widespread adoption of autonomous vehicles could lead to a shift from individual vehicle ownership to a system of shared-ride autonomous taxis (aTaxis) operated in fleets. This paper explores strategies for fleet management in large-scale systems, where all travel demand is served by aTaxis. Using New Jersey as a case study, this work looks at the implications that various empty vehicle repositioning strategies have on fleet size, empty vehicle miles traveled, and level of service provided.

Results show that repositioning vehicles locally can decrease the number of vehicles on the road significantly while still providing a high level of service. The maximum fleet size needed to operate the system and serve all passengers requires 3,232,096 vehicles and incurs a empty vehicle repositioning cost of 8.3% of the total vehicle miles traveled. The smallest feasible fleet size analyzed of 1,069,782 vehicles allows for 96.8% of passengers to be served with a wait time of at most 5 minutes beyond the advertised level of service. This method also incurs a empty vehicle repositioning cost of 5.2% of total vehicle miles traveled. These fleet sizes are a significant decrease from the 6,874,100 vehicles on the road in New Jersey in 2014.


INTRODUCTION

Autonomous vehicle technology has progressed significantly in recent years, with major investments from both car and technology companies. It is not too hard to envision that autonomous vehicles will soon be available to the public. The universal adoption of autonomous vehicles will bring massive changes to personal mobility in many ways. One of the key potential changes is a shift away from personal automobile ownership to a fleet service providing mobility on demand in driverless autonomous taxis (aTaxis). With aTaxi fleets, the challenge arises of operating and managing vehicles — deciding the number of vehicles to purchase and moving the vehicles between trips to best utilize the fleet.

This work builds upon Kornhauser and et al. (2016) [1], which analyzes ridesharing opportunities for New Jersey and takes a preliminary look at various repositioning strategies for managing an aTaxi fleet. Fagnant and Kockelman (2016) [2] simulate a system of aTaxis or Shared Autonomous Vehicles (SAVs) in Austin, Texas with dynamic ridesharing but at a low level of market penetration to determine the optimal fleet size. The trips taken in their simulation represent 1.3% of trips taken in the region. Pavone et al. (2012) [3] perform empty vehicle repositioning, or load rebalancing, on a simulated environment with randomly generated passengers.

This paper focuses on strategies for managing large-scale fleets of aTaxis in New Jersey and their effects on fleet size and level of service provided in scenarios where all travel demand is served by aTaxis. Given a constant demand, which is representative of actual travel demand in New Jersey, an upper and lower bound for the fleet size needed to operate the system is determined. Then, using two empty vehicle repositioning strategies for moving unused vehicles, the costs of reducing the fleet size from the upper bound is analyzed in terms of empty vehicle miles traveled and level of service provided. Unlike other case studies where empty vehicle repositioning strategies have been anticipatory (Fagnant and Kockelman (2015) [4]), that is, moving vehicles based on expected demand, this paper uses reactionary local repositioning strategies, moving vehicles after demand is known.

NEW JERSEY TRAVEL DEMAND

The data set used in this analysis is a set of generated trips representing the travel demand of New Jersey’s residents on a typical weekday. These trips were generated based on the population characteristics of New Jersey by Kornhauser and et al. (2012) [5]. Taking census data, a set of individuals with characteristics reflecting those of New Jersey’s residents was synthesized. Each individual was assigned to a home and work or school location. On an average day, an individual’s trips are assumed to start and end at home. Trips are assigned departure times based on probabilistic distributions of work and school schedules. Each trip is assumed to have 1 person and the complete set of trips generated contains about 30.5 million individual trips, with the average individual making 3.41 trips per day [5].

These trips were then assigned to the following modes of transport: walk, train, car/vehicle, airplane. A casual ridesharing analysis, meaning travelers were not incentivized for ridesharing, was conducted on the vehicle trips by Kornhauser and et al. (2014) [6] with the following restrictions:

- Maximum destinations of 3
- Maximum departure delay of 5 minutes from the time the first passenger arrives
- Maximum circuitry, or additional distance traveled by any passenger, of 20%

This reduced the 30,125,587 individual vehicle trips to 10,479,382 ridesharing trips. The analyses
in this paper are conducted on the ridesharing trips with the assumption that the second day’s travel
demand is the same as the first day’s. Although this is an unrealistic assumption, because the trips
are synthesized based on probabilistic distributions, generating a second day’s trips will look very
similar to the first day’s, unless fundamental assumptions about the model are changed.

System Operation
To reduce the computational complexity associated with using the exact latitude and longitude
coordinates of pickup and dropoff locations, the state of New Jersey was discretized into 0.5 x 0.5 mile pixels. Latitude and longitude coordinates were converted to pixel numbers using the following formula:

\[
X_{\text{coord}} = \text{floor}(108.907^\circ (\text{longitude} + 75.6)) \\
Y_{\text{coord}} = \text{floor}(138.2^\circ (\text{latitude} - 38.9))
\]

The pixelation of New Jersey can be seen in Figure 1a and the pixelation of Princeton, NJ can be seen in Figure 1b.

The aTaxi system operates as follows: an aTaxi stand is placed at the center of each pixel. All aTaxi trips originate from and end at the aTaxi stand. Passengers will walk or bike to the stand from where ever they are in the particular pixel. Each aTaxi stand has the capacity to hold and dispatch as many vehicles as necessary. For the purposes of this analysis, we assume that there is one aTaxi operator for the entire state. This means that regardless of where a vehicle originated from, it can be used to fulfill a trip at its destination pixel.

This system operates with 4 different vehicle sizes: 3 passenger, 6 passenger, 15 passenger, and 50 passenger. For simplicity, the number of passengers on the trip strictly dictates which vehicle they are assigned to, i.e. a trip with 5 people is always assigned to a 6 passenger vehicle, even if there are 3, 15, or 50 passenger vehicles available. For trips with more than 50 passengers, multiple 50 passenger vehicles are always used to fulfill the trip.

BASELINE REPOSITIONING METHODS
First, an upper and lower bound for the fleet size required to operate the system and serve all demand within the advertised level of service is established.

Lower Bound
To establish the lower bound, assume that vehicles can be moved infinitely fast when they drop off their final passengers. As soon as this happens, it is instantly moved to a pixel that needs a vehicle for a departure and is used there. Under this assumption, the fleet size needed is the maximum number of vehicles on the road at any given point during the day. Discretizing time by minutes, the minimum fleet size is the maximum vehicles on the road at any given minute during the day, shown in Figure 2. Note that for 15 and 50 passenger vehicles, the peak is much larger than the next largest number of vehicles on the road. This means that a nontrivial number of vehicles that must be purchased to satisfy this peak demand would be unused for the rest of the day.

This lower bound establishes the smallest fleet size necessary to serve all of the demand and provides a benchmark for comparison of the various empty vehicle repositioning strategies that will be used. However, in this case, because the assumption is made that for each departure, a vehicle instantly arrives from some other location with an available vehicle, the repositioning cost, or empty miles traveled, is unknown.
A “naive” repositioning strategy is used to establish the upper bound of vehicles needed. In the naive repositioning strategy, empty vehicles are moved only once per day. In this scenario, at the beginning of the day, there are enough vehicles at each pixel such that all trips are able to be served. At midnight, empty vehicles are moved so that the next day’s demand can be served without adding more vehicles. The repositioning of empty vehicles in the entire system is referred to as Early Morning Repositioning (EMR). While repositioning, the assumption is made that the vehicles are able to move infinitely fast to their destination pixels.

This strategy is implemented as follows. During the day, trips departing from a pixel are satisfied by either vehicle at the pixel or one that was brought to the pixel from the super source. Each arrival at a pixel increases the supply at the pixel and each departure decreases the supply. If there is a departure for which there is no available supply, a vehicle is brought from the super source. On day one, assume that vehicles can be brought to the pixel infinitely fast. For each subsequent day, the number of vehicles needed at each pixel at the beginning of the day is known.
(a) 3 passenger vehicles  
(b) 6 passenger vehicles  
(c) 15 passenger vehicles  
(d) 50 passenger vehicles

**FIGURE 2**: Number of vehicles on the road as a function of time. The minimum number of vehicles needed to operate the system is the maximum of each vehicle type.
since demand is the same, and the fleet can be positioned accordingly. The minimum fleet size
needed to operate the system is the number of vehicles that needed to be brought from the super
source during the day.

4 Early Morning Repositioning Cost
In repositioning the system, the goal is to minimize the total number of empty vehicle miles trav-
elled. This problem can be solved using the classic transportation linear program. The linear
program is modeled as follows:

\[
\begin{align*}
\min & \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} D_{ij} T_{ij} \\
\text{subject to} & \quad T \geq 0 \\
& \quad T_{ii} = 0 \\
& \quad \sum_{i \in \mathcal{I}} T_{ij} = A_j, \quad \forall j \in \mathcal{J} \\
& \quad \sum_{j \in \mathcal{J}} T_{ij} = P_i, \quad \forall i \in \mathcal{I}
\end{align*}
\]

(1)

where:

- \(D_{ij}\) is the distance between pixel \(i\) and pixel \(j\), calculated as \(1.2 \times D_{\text{cartesian}}\).
- \(T_{ij}\) is trip matrix, or the number of vehicles moved from pixel \(i\) to pixel \(j\).
- \(\mathcal{I}\) and \(\mathcal{J}\) are the set of active pixels.
- \(P_i\) is the number of excess vehicles available at the pixel \(i\).
- \(A_j\) is the number of vehicles needed at pixel \(j\).

Even though New Jersey has been discretized into pixels, solving the linear program over
all pixels still results in an extremely large optimization problem. For the smallest case, 50 passen-
er vehicles, there are about \(n = 8,000\) active pixels, which means there are roughly 6.5 million
variables to optimize over, as \(T\) is an \(n \times n\) matrix. To condense the problem, larger pixel blocks,
termed “super pixels,” are created, which are \(3 \times 3\) blocks of the standard \(0.5 \times 0.5\) mile pixels.
Starting at pixel \((0,0)\), pixels are grouped into blocks of 9 pixels. The super pixel containing pixel
\((0,0)\) has it center at pixel \((1,1)\). The super pixel center for any given pixel, \((x,y)\) is calculated as:

\(3 \times \text{floor}(x/3) + 1, 3 \times \text{floor}(y/3) + 1\)

The supply or demand at each super pixel is simply the sum of the supply or demand of
the smaller pixels within the super pixel. Distances between super pixels are calculated from the
center of each super pixel. This is able to greatly reduce the dimensionality of the linear program
in the largest case from \(n \approx 20,000\) to \(n \approx 3,000\), which can be solved in reasonable time using an
LP solver.

Though the distances that vehicles have to travel to and from the center of the super pixel
are not considered, there are both cases where vehicles travel both more than in a pixel to pixel
repositioning and less than in a pixel to pixel repositioning. These cases would average out, making
super pixel to super pixel repositioning a good approximation for pixel to pixel repositioning.

Results
A summary of the upper and lower bound fleet sizes for each vehicle size as well as the EMR
cost, is shown in Table 1. A minimum of 821,646 total vehicles and a maximum of 3,232,096
Zhu and Kornhauser

TABLE 1: Summary of baseline metrics

<table>
<thead>
<tr>
<th></th>
<th>3 Passenger</th>
<th>6 Passenger</th>
<th>15 Passenger</th>
<th>50 Passenger</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vehicle Trips Served</td>
<td>8,659,171</td>
<td>1,526,478</td>
<td>257,149</td>
<td>36,584</td>
<td>10,479,382</td>
</tr>
<tr>
<td>Lower Bound Fleet Size</td>
<td>632,947</td>
<td>133,275</td>
<td>43,957</td>
<td>11,467</td>
<td>821,646</td>
</tr>
<tr>
<td>Upper Bound Fleet Size</td>
<td>2,425,673</td>
<td>621,132</td>
<td>154,996</td>
<td>30,295</td>
<td>3,232,096</td>
</tr>
<tr>
<td>Upper Bound Empty Miles</td>
<td>16,039,770</td>
<td>5,515,587</td>
<td>854,912</td>
<td>204,821</td>
<td>22,615,090</td>
</tr>
</tbody>
</table>

Total vehicles are needed to operate the system. The EMR cost, compared to the loaded vehicle miles traveled (VMT), is 8.3%. Although this cost is low, there is a significant difference between the upper and lower bound of vehicles. With a low empty VMT, the number of vehicles needed is going to be the significant, controllable cost-driver of operating the system. To decrease the number of vehicles needed, additional repositioning strategies are implemented.

6 LOCAL REPOSITIONING

In the naive strategy, where cars are only moved empty at one point during the day, there are vehicles at pixels which are not being used for long periods during the day. To try and use vehicles more efficiently, local repositioning strategies is implemented, where vehicles are repositioned short distances during the day in order to increase fleet usage. First, since there is a departure delay of 5 minutes, a passenger who arrives at an aTaxi stand will wait for up to 5 minutes for the vehicle to depart. This means that for any particular departure, if there are no vehicles available at the pixel to serve the trip, one can be attempted to be sourced from a nearby pixel.

Two local repositioning strategies are analyzed. In the first strategy, a simple strategy, when a passenger arrives, only pixels with available vehicles that can reach the departure pixel before the departure time is considered. Figure 3 shows the pixels that are within 5 minutes of driving time from the departure pixel, using the assumption that vehicles travel at an average speed of 30 miles per hour. The departure pixel is in the center and the number in each pixel represents the amount of time needed to drive from that pixel to the departure pixel.

This strategy is implemented as follows. First, a supply array is initialized with an initial distribution of vehicles. Then, each trip, ordered by departure time, is analyzed. If there is an available vehicle at the departure pixel, it is used and the supply is decreased at the departure pixel and increased at the arrival pixel of that trip. If there is not an available vehicle at the departure pixel, then vehicle supplies at nearby pixels are checked, in order of time that it takes to reach the departure pixel. If an available vehicle is found at any of these pixels, then the supply is decreased at the pixel where the vehicle was found and increased at the destination pixel. If a vehicle was not found after looking at the pixels that are within 5 minutes’ travel time of the departure pixel, then that trip is not able to be served. However, leaving trips unserved is an unrealistic strategy.

An adjustment is made in the second strategy, an extended search strategy, to serve all trips.
For each trip departure, the simple search algorithm is carried out. However, if there is no available vehicle that can arrive in time, then the search extends to pixels farther and farther away until an available vehicle is found. This means that the passenger will have to wait longer than 5 minutes before they depart. The arrival time at the destination pixel is adjusted accordingly.

APPLICATION AND RESULTS
Various fleet sizes between the lower and upper bounds are analyzed on the data set to determine the effects of each strategy on repositioning costs and level-of-service provided. Five different fleet sizes are considered: 10%, 20%, 30%, 40%, and 50% in between the lower and upper bounds. The number of vehicles in each fleet size is shown in Table 2. An initial distribution is chosen to be proportional to the distribution of the naive strategy, rounding up to ensure integer values. This initial distribution is not the most optimal, but it is a good approximation, as choosing a random distribution would result in much higher costs [7].

<table>
<thead>
<tr>
<th></th>
<th>Lower</th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
<th>50%</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td># of vehicles</td>
<td>821,646</td>
<td>1,069,782</td>
<td>1,292,278</td>
<td>1,602,261</td>
<td>1,738,484</td>
<td>1,960,819</td>
<td>3,232,096</td>
</tr>
</tbody>
</table>

**Empty Vehicle Repositioning Costs**
Both local strategies also require that the fleet be repositioned at the end of the day, similar to the naive strategy. The cost of the local repositioning strategy then, in terms of empty miles, is the empty miles traveled in the local repositioning and the empty miles traveled in the naive repositioning. This cost for each strategy and fleet size is shown in Figure 4, broken down the contributions from local repositioning and EMR.

Though empty VMT was expected to be higher with the addition of local repositioning, this was not the case. Overall, the empty vehicle repositioning cost is very low and even with local
repositioning, in all cases, the total number of empty repositioning miles is less than that of the naive strategy, which was 8.3% of loaded VMT. As the fleet size increases, the EMR empty miles increase as well, since there are more vehicles that need to be repositioned. As the number of vehicles decreases, the amount of local empty miles increases. Since fewer vehicles are available, operators need to look farther from their pixel to find an available vehicle.

Local repositioning costs in all cases are a very small percentage of the total repositioning costs, indicating that during the day, vehicles tend to travel near areas with more departures. This means that this system should be able to provide a good level of service, as empty vehicles are easily accessible.

Level-of-Service Provided
In the naive strategy, all passengers are served within the advertised level of service, but at a cost of very large fleets. With smaller fleets, a high level of service is still able to be maintained. As seen in Table 3, a large majority, over 80%, of trips are able to be served within the advertised level of service (5 minute departure delay), even when reducing the fleet size to 10% between the lower and upper bound fleet sizes. The percentage of trips served increases as the fleet size increases, since more vehicles are available. However, this means that there are vehicles that are idle during the day that could be used more efficiently.

Of the strategies analyzed, the extended search is the most realistic operating strategy. If there were no vehicles available within 5 minutes, a fleet operator would not leave the trip unserved, but continue to look for an available vehicle as to not lose customers. An evaluation of the additional wait times for passengers in the extended search strategy, shown in Table 4, shows a very high level of service for even the smallest fleet size analyzed. The percentage of passengers able to be served within the advertised level of service is considerably higher than the percentage

**FIGURE 4**: Comparison of empty vehicle repositioning cost, as a percentage of loaded vehicle miles traveled, for fleets of varying sizes
TABLE 3: Percentage of trips served within advertised level of service in each repositioning strategy

<table>
<thead>
<tr>
<th>Fleet Size</th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
<th>50%</th>
<th>Naive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple Strategy</td>
<td>82.7</td>
<td>87.7</td>
<td>91.3</td>
<td>94.1</td>
<td>96.1</td>
<td>100</td>
</tr>
<tr>
<td>Extended Search</td>
<td>86.7</td>
<td>89.6</td>
<td>92.0</td>
<td>94.1</td>
<td>95.9</td>
<td>100</td>
</tr>
</tbody>
</table>

of trips able to be served within the advertised level of service in most cases. This indicates that a large part of the trips that are not able to be served within the advertised level of service are trips in smaller vehicles with fewer passengers.

TABLE 4: Percentage of passengers served as wait time increases beyond advertised level of service for various fleet sizes

<table>
<thead>
<tr>
<th>Fleet Size</th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
<th>50%</th>
<th>50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within advertised</td>
<td>95.4</td>
<td>96.4</td>
<td>97.2</td>
<td>97.9</td>
<td>98.6</td>
<td></td>
</tr>
<tr>
<td>Within advertised + 1 minute</td>
<td>95.7</td>
<td>96.3</td>
<td>97.5</td>
<td>98.2</td>
<td>98.8</td>
<td></td>
</tr>
<tr>
<td>Within advertised + 5 minutes</td>
<td>96.8</td>
<td>97.8</td>
<td>98.5</td>
<td>99.0</td>
<td>99.5</td>
<td></td>
</tr>
<tr>
<td>Within advertised + 10 minutes</td>
<td>97.8</td>
<td>98.7</td>
<td>99.2</td>
<td>99.6</td>
<td>99.8</td>
<td></td>
</tr>
</tbody>
</table>

These results have significant implications for the fleet size needed to operate the system. More than 98% of passengers are able to be served within the advertised level of service with a fleet of 50% between the minimum and maximum fleet sizes. Even with a fleet size of 10% between the minimum and maximum fleet sizes, 95.4% of passengers can be served within the advertised level of service and nearly 98% of passengers are served within an additional 10 minute wait. Given the extremely high level of service provided with the smaller fleets, it is both unnecessary and impractical to implement the naive strategy.

One drawback of the extended search strategy, as implemented, is that although most additional wait times are short, the distribution of wait times has a heavy right tail, with wait times of over an hour in a few extreme cases.

LIMITATIONS AND FUTURE WORK

This paper serves as a starting point for many future analyses of large-scale aTaxi fleet operations. A few are highlighted. Many simplifying assumptions were made in these analyses which can be adjusted for future work. For example, the analyses in this paper only consider a single state-wide aTaxi operator for all of New Jersey. In an operational scenario more comparable to the traditional taxi systems of today, taxis would be owned and operated regionally, with restrictions on pickups outside of a taxi’s home region. This type of operation would have a nontrivial effect on empty repositioning costs, fleet size, and efficiency that need to be further explored.

As mentioned previously, the number of passengers on each trip determines the vehicle...
size used. However, there may be scenarios where it is more efficient to use multiple vehicles of smaller sizes or larger vehicles to serve the trip. This could also change depending on the ridesharing strategies used. In this data set, the ridesharing was implemented with inefficiencies, leading to large vehicles traveling large distances with only a few passengers [7].

Finally, these analyses can be used to determine an optimal fleet size to operate a statewide system, given costs of purchasing and operating vehicles, as well as penalties for passengers not served within the advertised level of service.

8 CONCLUSION
9 This paper analyzed the effects of empty vehicle repositioning strategies on fleet size and level-of-service provided in a large-scale, mobility-on-demand aTaxi system in New Jersey without the need to anticipate future demand. The scenarios analyzed suggest that, with local repositioning, all of the travel demand of New Jersey can be served with a fleet of shared-ride aTaxis that is much smaller than the current fleet of vehicles operating in New Jersey. In 2014, there were 6,874,100 registered motor vehicles in New Jersey, including public buses [8]. The demand can be well served with an aTaxi fleet 10% between the lower and upper bound, or 1,069,782 vehicles. With this fleet size, one aTaxi would replace more about 6 traditional vehicles. This would have tremendous benefits in terms of decreasing environmental pollution and vehicular congestion, as well as decreasing the need for parking structures.

While full adoption of autonomous vehicles as the primary mode of transport is still far off, the benefits of large-scale aTaxi systems can be quantified and can play a major role in reducing the congestion and environmental issues that are faced in many metropolitan areas today.
REFERENCES


