Short-Term Repositioning for Empty Vehicles on Ride-Sourcing Platforms

Hai Wang
Carnegie Mellon University, haiwang@cmu.edu
Singapore Management University, haiwang@smu.edu.sg

Zhengli Wang
Stanford University, wzl@stanford.edu

Motivation Ride sourcing companies, such as Uber, Lyft, and Didi, have been able to leverage on internet-based platforms to connect passengers and drivers. These platforms facilitate passengers and drivers' mobility data on smartphones in real time, which enables a convenient matching between demand and supply. The imbalance of demand (i.e., passenger requests) and supply (i.e., drivers) on the platforms causes many unserved passenger requests and empty vehicles with idle drivers to exist at the same time, which poses a challenging problem for the platform. To address these challenges, some platforms display heat maps of surge-pricing multipliers or real-time demand to drivers, and anticipate that such information will induce more idle drivers to self-reposition by cruising to regions with high demand and/or low supply. Some platforms attempt to provide direct repositoning guidance to drivers, by suggesting that drivers cruise to a specific region. In practical terms, drivers are more likely to follow guidance that directs them to a region close to their current location. Therefore, in reality, platforms divide the entire service area into many hexagonal regions, and usually provide repositioning guidance for a driver to relocate from his/her current hexagonal region to an adjacent region. Such repositioning suggestions are normally provided regularly to idle drivers at short intervals (e.g., every 5 minutes). Great effort has been devoted to solving the repositioning problem; see Wallar et al. (2018); Yu et al. (2019); and Braverman et al. (2019). In this paper, we focus on short-term repositioning policy for empty vehicles with idle drivers on ride-sourcing platforms. We optimize repositioning guidance that directs drivers to relocate to adjacent regions for a short period. Compared with previous work, we make the following contributions:

— We propose a two-stage framework to study short-term repositioning policy;

— We provide and prove a closed-form optimal repositioning solution if all regions are connected with unit travel time (stage 1 model);

— We provide an optimization formulation for any actual connective topology of regions, and prove some performance bounds for the actual feasible repositioning (stage 2 model); and

—We describe and evaluate the performance of the two-stage framework in a set of numerical experiments using actual data from a ride-sourcing platform.

Problem Statement We construct a graph by dividing the city into K hexagonal regions (i.e., K nodes). Adjacent regions have connecting edges (i.e., M edges). We can obtain two vertex-edge matrices A and O, where A is a $K \times M$ vertex-edge matrix that indicates origin and destination, and O is a transformed version of A given by $O = max(A^-, 0)$. We divide a day into consecutive time periods $t = 1, 2, ...$, where each time period indicates a short-term interval of 5 or 10 minutes. From historical data, we can estimate the demand distribution of requests at each node i at each time t, denoted as $C_{i,t}$. At any time t for each region i, the platform has the following information: the number of empty vehicles with idle drivers $u_{i,t}$, the number of unserved passenger requests $p_{i,t}$, and the maximum number of vehicles $\eta_{j,t}$ allowed to travel across edge j. Moreover, the platform at time t can also infer the number of vehicles $d_{i,t+1}$ that will become empty (and hence available) at each region for the next time period $t+1$ based on their current service status. Then, for a short-term repositioning at each time t, the platform's decision is to reposition a flow on each edge to achieve maximum expected matches in time $t+1$. The problem can be formally formulated as

$$\begin{aligned} \max_{\vec{f_t}} \ & E\sum_{i=1}^{K} \min\{p_{i,t} + C_{i,t}, d_{i,t+1} + n_{i,t+1}\} \\ s.t. \ & o_i \vec{f_t} \le u_{i,t}, \forall i \in \{1, 2, ..., K\} \\ & a_i \vec{f_t} = n_{i,t+1}, \forall i \in \{1, 2, ..., K\} \\ & 0 \le f_{j,t} \le \eta_{j,t}, \forall j \in \{1, 2, ..., M\} \end{aligned}$$

Approach The formulated problem is a convex optimization problem that is hard to solve due to the presence of expectation at every node and the large number of constraints and decision variables. We tackle the difficulty by proposing a two-stage framework that considers regions' connective topology under the city's road map, real-time empty and occupied vehicles' locations, and historical and instant passenger requests. Specifically, **in the first stage (stage 1)**, we solve the above optimization problem assuming that the graph is complete— i.e., all regions are connected with unit travel time. The optimal solution is the target we want to achieve. **In the second stage (stage 2)**, we take into account the graph structure of the actual connective topology of regions, and direct the repositioning flow so that in the next time period, each region will obtain vehicles as close as possible to the target (i.e., optimal solution from stage 1). The advantages are that in stage 1 we can obtain closed-form optimal solution as a target that can be computed fast; in stage 2, we can use existing network flow algorithms to obtain a feasible solution considering actual topology.

In stage 1, we assume a complete graph and solve the problem

$$\begin{aligned} \max_{\{n_{i,t+1}\}_{i=1}^{K}} \quad & E\sum_{i=1}^{K}\min\{p_{i,t}+C_{i,t}, d_{i,t+1}+n_{i,t+1}\} \\ s.t. \quad & \sum_{i=1}^{K} n_{i,t+1} = \sum_{i=1}^{K} u_{i,t} \\ & n_{i,t+1} \geq 0, \forall i \in \{1,2,...,K\} \end{aligned} \tag{1}$$

The optimal solution is given in a closed form as the *CDF Balance Law* in Theorem 1.

THEOREM 1. *The optimal number of vehicles to be repositioned to each region i for the next time period, $n^*_{i,t+1}$, is given by*

$$\begin{aligned} n^*_{i,t+1} &= F_i^{-1}(w) - d_{i,t+1} \ , \ \textit{if } F_i^{-1}(w) > d_{i,t+1} \\ n^*_{i,t+1} &= 0 \ , \ \textit{if } F_i^{-1}(w) \leq d_{i,t+1} \end{aligned} \tag{2}$$

for a certain value w, where F_i is the CDF for the demand distribution at region i.

In stage 2, we set the optimal solutions $n^*_{i,t+1}$ from stage 1 as the target, and direct the repositioning flow so that vehicles as close as possible to the target by imposing a L^2 penalty:

$$\begin{aligned} & \min_{\vec{f_t}} (A\vec{f_t} - \vec{n}_{t+1})'(A\vec{f_t} - \vec{n}_{t+1}) \\ & s.t. \ o_i\vec{f_t} \leq u_{i,t}, \forall i \in \{1,2,...,K\} \\ & \quad 0 \leq f_{j,t} \leq \eta_{j,t}, \forall j \in \{1,2,...,M\} \end{aligned} \tag{3}$$

We then provide a lower bound on the optimal value in stage 2 described in Theorem 2.

THEOREM 2. *A lower bound for the optimal objective value in stage 2 formulation (3) is*

$$||\vec{n}_{t+1}||^2 - 2(\vec{n}'_{t+1}A)^+ \min\{\vec{\eta_t}, \vec{u}_t^{(E)}\} \tag{4}$$

where $u_t^{(E)} = \{u_{j,t}^{(E)}\}_{j=1}^{M}$, $u_{j,t}^{(E)} = u_{i,t}$ if edge j starts from node i and $\min\{\vec{\eta_t}, \vec{u}_t^{(E)}\}$ represents component-wise minimum of the two vectors $\vec{\eta_t}$ and $\vec{u}_t^{(E)}$. In practice, we can use the lower bound as a heuristic measure of how good the decision is.

Demand	Low (12 pm to 1 pm)		Intermediate (2 pm to 3 pm)		High (8 am to 9 am)	
Number of Requests	21,415		26,286		33,660	
Max. Allowed Pickup Time	5 min	10 min	5 min	10 min	5 min	10 min
Stay and Wait	27.2%	28.8%	22.3%	24.2%	20.1%	21.2%
Random Walk	45.0%	48.1%	37.9%	40.4%	33.7%	34.9%
Target Cruise	74.2%	88.9%	66.4%	81.4%	59.4%	74.7%
Short-term Repositioning	92.3%	93.1%	87.4%	91.0%	78.1%	83.5%

Table 1 Service Level (proportion of served requests) for fleet size 8,000 in Numerical Experiments

Numerical Results We extracted real data from Didi Chuxing, the largest on-demand ride-sourcing platform in China, and conducted a set of numerical experiments on the proposed short-term repositioning policy. Table 1 reports the fleet size, number of requests, and maximum allowed pickup time for three demand scenarios in the experiments. We compare the performance of our policy, assuming that all vehicles follow the short-term repositioning guidance, with three benchmarks in terms of service level (i.e., proportion of served requests). The benchmark cases are (1) vehicles always stay and wait in the drop-off location of the previous order (i.e., stay and wait); (2) vehicles randomly cruise the city (i.e., random walk); and (3) vehicles always cruise to a nearby region with highest historical demand (i.e., target cruise). It is seen that the proposed short-term repositioning can increase the service level significantly compared with all benchmarks. Specifically, it increases the service level by around 5% to 10% above that of the target cruise. We also evaluate the impacts of order dispatch priority in Figure 1. The horizontal axis is the percentage of system-repositioning vehicles (i.e., the percentage of collaborative drivers who follow the short-term repositioning guidance; the rest are self-repositioning noncollaborative drivers using the target cruise strategy). The vertical axis is the service level. If the platform gives order-dispatch priority to self-repositioning vehicles, there are more empty system-repositioning vehicles on the platform—and hence more optimization space for future repositioning—which results in a higher service level. If the platform gives order-dispatch priority to system-repositioning vehicles, there are fewer empty system-repositioning vehicles, which results in a lower service level.

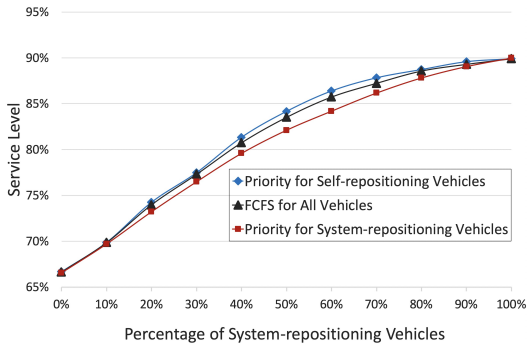


Figure 1 Impacts of Orde-dispatch Priority

References

Braverman A, Dai JG, Liu X, Ying L, 2019 *Empty-car routing in ridesharing systems. Operations Research* 67(5):1437–1452.

Wallar A, Van Der Zee M, Alonso-Mora J, Rus D, 2018 *Vehicle rebalancing for mobility-on-demand systems with ride-sharing. 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 4539–4546 (IEEE).

Yu X, Gao S, Hu X, Park H, 2019 *A markov decision process approach to vacant taxi routing with e-hailing. Transportation Research Part B: Methodological* 121:114–134.