Dynamic Inventory Relocation for a One-way Electric Car Sharing System with Uncertain Demand

Rui Liu
Chongqing University, School of Economics and Business Administration, liuruicqdx@cqu.edu.cn,

Shu Zhang
Chongqing University, School of Economics and Business Administration, zhangshu@cqu.edu.cn,

Ping Chen
Nankai University, Business School, chenpingteda@nankai.edu.cn,

1. Introduction and Problem Formulation

While the increasing ownership of private gasoline vehicles has caused issues such as air pollution, traffic congestion, and shortage of parking spaces, vehicle sharing systems have been promoted in many cities across the world to alleviate negative impacts from the adoption of private vehicles. In this study, we investigate the inventory relocation operation of a car-sharing system with electric vehicles (EV). We consider a one-way car sharing system. That is, the users can pick-up and return vehicles at different stations. Given that the pick-up and return demands from and to a station are uncertain, the car-sharing operating company may need to relocate vehicles between stations, in order to maximize the expected demand satisfied by the system. We assume that the car-sharing system has a homogeneous fleet of vehicles and there are several personnel hired for EV relocation. We assume that customers will not rent a vehicle with low battery and the energy consumptions of rented EVs are uncertain. Upon arrival at a station, the person working for EV relocation needs to decide which vehicle to choose and to which station to move the vehicle.

Given the stochasticities in customer demands and battery level of returned EVs, we formulate the problem as a Markov decision process. We define the problem on a complete graph (G, E), where $G(|G| = n + 1)$ corresponds to the depot 0 where the relocation personnel depart from and n stations where EVs are parked. The set E corrresponds to the set of edges associated with each pair of locations in set G. We represent by $N(N = G \setminus \{0\})$ the set of n stations. Let d_{ij} and t_{ij} represent the distance and deterministic travel time between stations i and j. We assume that each electric vehicle will consume $\delta \cdot d_{ij}$ units of energy when traveling on arc (i, j), where δ is the

energy consumption rate per unit of distance. Let V represent the set of EVs in the system. We assume that each EV has a battery capacity of Q. Let M be the set of personnel working on EV relocation (named as workers hereafter). Let P_i represent the total number of parking spaces at station $i \in N$ and $P = \max\{P_i\}_{i\in N}$. We assume that each parking space has a charging outlet. The customers are required to charge the EVs after they returned the vehicles to the parking spaces. We consider a discrete time-planning horizon $0, 1, \ldots, T$, where T is the point of time when the workers will be off work. We assume that all workers start working at the beginning of the time horizon. We consider one minute as the minimum time increment in the model.

The state of the system consists of information on the workers, EVs and parking space availabilities at stations. Specifically, let $(t, w) = (t_m, w_m)_{m\in M}$ be the vector represent information on workers, where $w_m (m \in M)$ is worker m's destination station (or the station the worker has just arrived) and $t_m (m \in M)$ is the time when the worker will arrive at station w_m (or the time she has just arrived at the station). If the workers are at the depot, we have $t_m = ?, \forall m \in M$. Let $(l, \mu) = (l_c, \mu_c)_{c\in V}$ be the vector represent information on EV $c \in V$, where l_c is the current location of the EV and $\mu_c \in [0, Q]$ is the current charge level of the vehicle. If EV $c \in V$ is at one of the stations, we have $l_c \in N \setminus \{0\}$. If the EV is rented by a customer, we are unaware of the vehicle's current location and set $l_c = ?$. If EV c is at a station or being relocated by a worker, we can observe the value of μ_c. If the EV is rented by a customer, we are unaware of the vehicle's current charge level and set $\mu_c = ?$. Let $q = (q_i)_{i\in N} (q_i \in [0, P_i])$ represent the available parking spaces at station i. Thus, the state of the system is represented by $s_k = ((t, w), (l, \mu), q)$. The state space S is defined on $[0, T]^{|M|} \times \{N \cup \{?\}\}^{|M|} \times \{N \cup \{?\}\}^{|V|} \times \{[0, Q] \cup \{?\}\}^{|V|} \times [0, P]^{|N|}$.

A decision needs to be made when the workers are at the depot. After leaving the depot, a decision epoch is triggered by the arrival of one or more workers at their destinations. Let ϕ_k represent the time of the k-th decision epoch. Thus, we have $\phi_k = \min_{m\in M} t_m$. Let $\check{m} = \arg\min_{m\in M} t_m$ represent the worker who has just arrived at her destination station. At station $w_{\check{m}}$, worker $\check{m}$ can either move an EV to another station, or leave for another station without relocating a vehicle. The latter is applicable when there are no EVs available to be relocated (either all EVs are rented before the worker arrives, or the EVs' charge levels are not sufficient to cover the distance to another station). Let $a_k = (\check{c}, \psi) (\check{c} \in V, \psi \in N)$ represent the action made at state s_k. We note that if the worker does not choose any EV, then $\check{c} \in \{\emptyset\}$. If deciding to relocate a vehicle, the worker can only choose a station which has available parking space and choose an EV whose charge level is sufficient to cover the distance to that station. Thus, the action space for state s_k is $A(s_k) = \left\{(\check{c}, \psi) \in \left\{V \times N\right\} : l_{\check{c}} = w_{\check{m}}, \mu_{\check{c}} \geq \delta d_{w_{\check{m}}\psi}, q_\psi \geq 1\right\} \cup \left\{(\check{c}, \psi) \in \left\{\{\emptyset\} \times N\right\}\right\}$.

Let Z_{it}^c be a binary random variable denoting whether vehicle c is requested by a customer at station i at time t and z_{it}^c is a realization of Z_{it}^c. If $z_{it}^c = 1$, station i received a rental requests for

vehicle c at station i. Otherwise, $z_{it}^c = 0$. The reward collected in the transition from state s_k to s_{k+1} is the total demand satisfied between periods ϕ_k and ϕ_{k+1}. Let $R_k(s_k, a_k)$ be the reward function given the current state s_k and action a_k. Then we have $R_k(s_k, a_k) = \sum_{i \in N} \sum_{t=\phi_k}^{\phi_{k+1}} \min\{z_{it}^c, h_{it}^c\}$. Let Π be the set of all Markovian decision policies. We seek a policy $\pi \in \Pi$ that maximizes the total expected reward $E\big[\sum_{k=1}^{K} R_k(s_k, \rho_{s_k}^{\pi}(s_k))|s_0\big]$, where $\rho_{s_k}^{\pi}(s_k) : s_k \mapsto A(s_k)$ is a function specifying the action to select at decision epoch k while following policy π.

2. Solution Approach

Due to the curse of dimensionality presented in this model, we are not able to find an optimal policy to the MDP using backward dynamic programming. Thus, we use an approximate dynamic programming algorithm to solve this problem. Specifically, we implement an approximate value iteration (AVI) algorithm. We refer the readers to Powell (2011) for an overview of the AVI methods. Given the large size of our state space, however, we cannot operate on the original state space. We aggregate the state space by the information on time and the average available number of EVs over all stations. Specifically, we represent workers' information (t, w) by $\phi_k = \min\{t_m\}_{m \in M}$, because ϕ_k is the time when a state transition occurs. We divide the whole area where EV stations are located into sub-areas, each containing one or more stations. Let $\mathfrak{I}$ represent the set of subareas. Let $\nu_i (i \in \mathfrak{I})$ be the number of stations in subarea i and θ_i be the available EVs in this area. We assume that an EV is available for renting only if its charge level is above a threshold level $\breve{\mu}$. We represent by $\omega_i = \frac{\theta_i}{\nu_i}$ the average number of available EVs per station in area $i \in \mathfrak{I}$. Then we use aggregation $\mathfrak{G}(s_k)$ to represent the original state space by a $1 + |\mathfrak{I}|$ dimension vector $\mathfrak{G}(s_k) = \varrho_k = (\phi_k, (\omega_i)_{i \in \mathfrak{I}})$.

In the AVI, a lookup table (LT) is used to store the values of the states. We represent by $\mathfrak{L}$ the LT for vector ϱ_k. According to the above discussion, we need a $1 + |\mathfrak{I}|$ dimensional LT. If we assume the minimal time increment as one minute and a one-by-one representation of station numbers, the cardinality of LT $\mathfrak{L}$ is still high after state space aggregation. As we may not be able to have enough observations for each entry in LT $\mathfrak{L}$, we considered a partitioned LT $\mathfrak{C}$, where $\mathfrak{C}$ is defined by a partitioning $\mathscr{P} : \mathfrak{L} \to \mathfrak{C}$ that groups vectors $\varrho_1, \ldots, \varrho_q \in \mathfrak{L}$ to $\eta = \mathscr{P}(\varrho_1) = \ldots = \mathscr{P}(\varrho_p)$. In this study, we adopt the dynamic lookup table (DLT) as in Ulmer, Mattfeld, and Köster (2017). Instead of using equivalent interval lengths of the parameters, we start with large intervals and sequentially decrease the interval length based the number of entries observed in the interval. In the preliminary experiments, we use a fixed threshold entry number κ. For each vector in $\mathfrak{C}$, if the observed number of entries reaches κ, we divide the current interval length by half.

3. Experimental Design and Preliminary Results

In the preliminary experiments, we consider an instance with 15 stations, each having seven empty parking spaces and five fully charged EVs at the beginning the horizon. The stations are

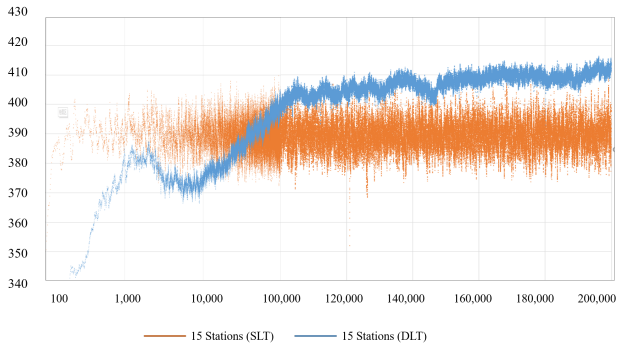


Figure 1 **Preliminary Results**

randomly located in four subareas. We consider that there are two workers working for 8 hours each day for EV relocation. We assume that a fully charge EV is of 100 units of energy and traveling for one unit of distance will consume 0.5 unit of energy. We assume that it takes 50 minutes to fully charge an EV via the fast charging mode. Between two decision epochs, we assume that the chances that none, one, two, three, and four EVs are rented for a station are 70%, 20%, 5%, 4%, and 1%, respectively. We set the threshold level $\check{\mu} = 20$, which indicates that EVs whose charge level is above 20 units of energy is available to be rented. We set the initial intervals for ϕ_k in the DLT as [0, 59], [60, 119] in minutes and that for $\omega_i (i \in \mathfrak{I})$ as [0, 3] and [4, 7]. We set the threshold value $\kappa = 10,000$, which indicates that after κ iterations, we will split the current interval into half. We compare the effectiveness of the DLT with a static lookup table (SLT), in which we use fixed interval lengths (the same as the original interval lengths in the DLT) through all iterations. We run 200,000 iterations for the AVI with both lookup tables.

In Figure 1, we present the preliminary computational results. The orange and blue lines correspond to the expected objectives obtained by the SLT and DLT, respectively. As shown in the figure, with the STL the expected objective converges earlier than with the DLT, but the variances of the expected objectives are much greater than that from using the DLT. The average expected objective for the last 5000 runs with the SLT and the DLT are 393.6 and 411.6, respectively. From the preliminary experiments, we learn that the advantage that the DLT has over the SLT is not significant. Thus, the next step is to improve the mechanism in updating the DLT, e.g., when and how to decrease the interval length in the lookup table.

References

Powell WB, 2011 *Approximate Dynamic Programming: Solving the Curses of Dimensionality, Second Edition* (NJ, USA: John Wiley and Sons, Hoboken).

Ulmer MW, Mattfeld DC, Köster F, 2017 *Budgeting time for dynamic vehicle routing with stochastic customer requests. Transportation Science* 52(1):20–37.