Online operation strategies for crowdsourced mobile charging service

A. Yiming Yan[a], B. Xi Lin[b], C. Yitong Yu[a] and David Z.W. Wang [a,*]

[a] <School of Civil and Environmental Engineering>, <Nanyang Technological University>, <Singapore>
a.yiming007@e.ntu.edu.sg, b.x-lin@mail.tsinghua.edu.cn
[b] <Department of Civil and Environmental Engineering>, <University of Michigan>, <Ann Arbor>, <United States>
c.yitong001@e.ntu.edu.sg
d.wangzhiwei@ntu.edu.sg
[*] Corresponding author

Extended abstract submitted for presentation at the Conference in Emerging Technologies in Transportation Systems (TRC-30)
September 02-03, 2024, Crete, Greece

April 28, 2024

Keywords: (Crowdsourced mobile charging; Charging service for Electric vehicles; Reinforcement learning; Online operation)

1 Introduction

Electric Vehicles (EVs) are widely recognized as a promising solution for reducing emissions and saving energy. There has been a fast-growing adoption of EVs in recent years, mobilizing the transportation landscape towards a more environmentally friendly future (Ferrero *et al.*, 2016). However, despite considerable technological and market progress, charging EVs is still less convenient than conventional refueling vehicles. The abovementioned issue calls for an integrated and efficient EV charging system to cater to massive and diverse charging demands in the future electric mobility system. Based on charging location flexibility, existing recharging modes for EVs can generally be classified into fixed charging and mobile charging. This study considers a mobile charging crowdsourcing online platform that utitlizes crowdsourced vehicles to provide mobile charging services for EV users. The primary objective of this study is to determine the optimal operation strategies, such as the matching between mobile charging demand and supply, repositioning of the mobile chargers, so as to maximize the profit gained by the operation of crowdsourced mobile charging service.

Let us consider a multi-zone service area. The area contains a set of pre-determined zones, denoted by $\mathcal{Z}$. Each zone contains some potential customer locations (e.g., parking lots), where EVs could be parked for mobile charging services. In zone $z \in \mathcal{Z}$, the set of potential customer locations is denoted by $\mathcal{P}_z$, and $\mathcal{P} \triangleq \bigcup_{z\in\mathcal{Z}} \mathcal{P}_z$. There are some recharging facilities for the mobile chargers to get recharged (or get their batteries swapped), and the set of all recharging facilities is denoted by $\mathcal{R}$. There could exist some zones without any recharging facility. The travel time between two locations $i, j \in \mathcal{P} \bigcup \mathcal{R}$ is a unchanged magnitude, denoted as $\bar{t}_{ij}$; $\bar{t}_{ii} = 0$ for any $i \in \mathcal{P} \bigcup \mathcal{R}$.

A platform is operating the mobile charging service in this area, and the operation is a dynamic process with certain level of stochasticity. At any time, there could be a newly generated recharging order at certain potential customer location, with an associated latest allowable service completing time (indicating the expected time that the vehicle leave the parking spot) and the associated charging amount. Each order is with a recharging earning for the platform that is directly related to the charging amount. If a vehicle does not get a complete recharge before its latest allowable time, the order is canceled, and there will be a penalty associated with the platform operation. On the other hand, the platform is operating a flexible fleet of crowdsourced mobile chargers, such that there could be mobile chargers joining or leaving the fleet at any time. Each mobile charger is with a maximum battery level and a safety battery level; if the battery level is below the safety threshold after serving one order, the charger needs to go to one recharging facility. We ignore the energy consumption associated with traveling of mobile chargers. The goal of this study is to design a methodological framework for operating the mobile charger fleet, including routing, recharging, and repositioning among zones.

2 Methodology

The methodological framework for the online operation of the crowdsourced mobile charging services is designed as a two-layer structure, as shown in Figure 1. The lower layer concerns the routing (or equivalently, order assignment) of mobile chargers within each zone, and the decision-making frequency is relatively high (e.g., every 15 minutes); and the upper layer decides the inter-zone charger repositioning, with a relatively larger decision-making interval (e.g., every 45 minutes). The motivation for this two-layer framework is to balance the solution quality and computational loads. On one hand, by restricting the routing problem to each zone, the lower-layer problem can greatly alleviate the real-time computational burden; on the other hand, the upper-level repositioning decision can coordinate the routing of each zone in order for handling the spatio-temporal heterogeneity between supplies and demands, therefore achieving area-wise operating efficiency.

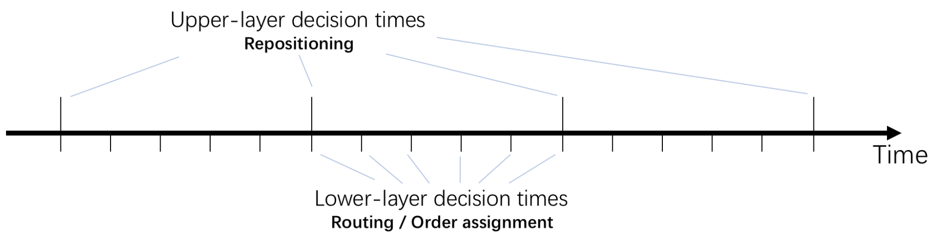

Figure 1 – *Illustration of the two-layer modeling framework.*

2.1 The lower-layer problem

The lower layer problem determines the routing (i.e., order assignment) of each mobile charger for each zone. Consider a specific zone at certain lower-layer decision time t_{d}, and for notational brevity we ignore the subscript of the zone and the time. The set of all mobile chargers in this zone at time t_{d} is denoted as $\mathcal{M}$. For a charger $k \in \mathcal{M}$, its earliest available time is $t_{\mathrm{e},k} \geq t_{\mathrm{d}}$, suggesting the time that it completes the current order service or completes recharging; at $t_{\mathrm{e},k}$, its location is $L_{\mathrm{e},k} \in \mathcal{P} \bigcup \mathcal{R}$, the remaining battery level is denoted as $B_{\mathrm{e},k}$. The safety battery level of charger k is $\underline{B}_k$. On the demand side, the set of all unserved orders is denoted as $\mathcal{O}$. For an order $l \in \mathcal{O}$, its location is $\widehat{L}_l \in \mathcal{P}$, the amount of charge required is $\widehat{E}_l$, the charging time required is $\widehat{G}_l$, the charging revenue is $\widehat{Y}_l$, and the associated latest allowable service completing time is $\widehat{T}_l$. Besides, if a vehicle does not get a complete recharge before its latest allowable time, the order is canceled, and there will be a penalty $\hat{P}_l$ which is same for all orders no matter what the amount of charge required.

In real-world operation, the size of unserved orders could be very large. Also, as the operation is inevitably associated with certain level of stochasticity, it is of limited value to determine the routing of all mobile chargers throughout the whole day in one decision time. Thus, for the lower layer decision problem, we set the decision horizon to be $[t_\mathrm{d}, t_\mathrm{d} + H_\mathrm{d}]$, where H_d is the length of the decision horizon, and H_d should be larger than the lower-layer decision-making interval. We define a restricted order set $\widehat{\mathcal{O}}$ as a subset of $\mathcal{O}$ that contains all orders with the latest allowable service completing time earlier than $t_\mathrm{d} + H_\mathrm{d}$; if $\left|\widehat{\mathcal{O}}\right| < |\mathcal{M}|$, then we define $\widehat{\mathcal{O}}$ as the set containing $|\mathcal{M}|$ orders with the earliest allowable service completing times, and in this case the penalty does not exist. The goal of the decision problem at the current time is to assign the orders in $\widehat{\mathcal{O}}$ to the mobile chargers for maximizing the total payment to the platform.

With the above preparation, the lower-layer decision model in each zone is then formulated as follows.

$$\max_{\mathbf{x},\mathbf{u},\hat{\mathbf{t}},\hat{\mathbf{e}}} \sum_{n\in\{1,2,\ldots,\overline{N}\}} \sum_{k\in\mathcal{M}} \sum_{l\in\widehat{\mathcal{O}}} \widehat{Y}_l x_{kl}^n - \sum_{l\in\widehat{\mathcal{O}}} \hat{P}_l (1 - \sum_{n\in\{1,2,\ldots,\overline{N}\}} \sum_{k\in\mathcal{M}} x_{kl}^n) \tag{1}$$

$$\text{s.t.} \quad \sum_{n\in\{1,2,\ldots,\overline{N}\}} \sum_{k\in\mathcal{M}} x_{kl}^n \leq 1 \quad \forall l \in \widehat{\mathcal{O}} \tag{2}$$

$$1 \geq \sum_{l\in\widehat{\mathcal{O}}} x_{kl}^n \geq \sum_{l\in\widehat{\mathcal{O}}} x_{kl}^{n+1} \quad \forall k \in \mathcal{M}, n \in \{1,2,\ldots,\overline{N}-1\} \tag{3}$$

$$u_{l,l'}^{k,n} \geq x_{kl}^n + x_{kl'}^{n+1} - 1 \quad \forall k \in \mathcal{M};\ l,l' \in \widehat{\mathcal{O}}, n \in \{1,2,\ldots,\overline{N}-1\} \tag{4}$$

$$u_{l,l'}^{k,n} \leq \frac{1}{2}\left(x_{kl}^n + x_{kl'}^{n+1}\right) \quad \forall k \in \mathcal{M};\ l,l' \in \widehat{\mathcal{O}}, n \in \{1,2,\ldots,\overline{N}-1\} \tag{5}$$

$$\hat{t}_k^1 = t_{\mathrm{e},k} + \sum_{l\in\widehat{\mathcal{O}}} x_{kl}^1 \left(\bar{t}_{kl} + \widehat{G}_l\right) \quad \forall k \in \mathcal{M} \tag{6}$$

$$\hat{t}_k^n = \hat{t}_k^{n-1} + \sum_{l,l'\in\widehat{\mathcal{O}}} u_{l,l'}^{k,n-1} \left(\bar{t}_{l,l'} + \widehat{G}_{l'}\right) \quad \forall k \in \mathcal{M}, n \in \{2,\ldots,\overline{N}\} \tag{7}$$

$$\hat{t}_k^n \leq \widehat{T}_l + M(1 - x_{kl}^n) \quad \forall k \in \mathcal{M}, l \in \widehat{\mathcal{O}}, n \in \{1,\ldots,\overline{N}\} \tag{8}$$

$$\hat{e}_k^1 = B_{\mathrm{e},k} \quad \forall k \in \mathcal{M} \tag{9}$$

$$\hat{e}_k^n = \hat{e}_k^{n-1} - \sum_{l\in\widehat{\mathcal{O}}} x_{kl}^{n-1} \widehat{E}_l \quad \forall k \in \mathcal{M}, n \in \{2,\ldots,\overline{N}\} \tag{10}$$

$$\sum_{l\in\widehat{\mathcal{O}}} x_{kl}^n \leq \frac{\hat{e}_k^n}{\underline{B}_k} \quad \forall k \in \mathcal{M}, n \in \{2,\ldots,\overline{N}\} \tag{11}$$

$$x_{kl}^n, u_{l,l'}^{k,n} \in \{0,1\} \quad \forall k, l, l', n \tag{12}$$

The above problem contains four types of decision variables: x_{kl}^n is a 0-1 variable indicating whether charger k serves order l on its n^{th} place in the sequence; $u_{l,l'}^{k,n}$ is a 0-1 variable indicating whether charger k serves order l on its n^{th} place and then serves order l' on its $(n+1)^{\mathrm{th}}$ place in the sequence; $\hat{t}_k^n$ represents the service completing time for charger k on its n^{th} place in the sequence; and $\hat{e}_k^n$ represents the remaining battery level for charger k before serving its n^{th} order in the sequence. The objective function Eq.(1) aims at maximizing the total profit minus the penalty of unserved demand in the decision horizon, where $\overline{N}$ is the preset maximum number of orders assigned to each mobile charger. Eq.(2) states that each order is served by at most one mobile charger once. Eq.(3) states that a mobile charger serves at most one order in one place in the sequence, and meanwhile suggests the sequence feasibility. Eqs.(4)-(5) demand that: $u_{l,l'}^{k,n} = 1$ if and only if $x_{kl}^n = x_{kl}^{n+1} = 1$. Eqs.(6)-(7) state the updating rules for variable $\hat{t}_k^n$, and Eq.(8) requires that: if charger k serves order l on its n^{th} place in the sequence, the service completing time must be no later than the associated latest allowable time $\widehat{T}_l$. Similarly, Eqs.(9)-(10) state the updating rules for variable $\hat{e}_k^n$, and Eq.(11) requires that a charger k can serve some order only if the remaining battery level is no less than the safety threshold $\underline{B}_k$. Finally, Eq.(12) states the binary nature of variables $x_{kl}^n, u_{l,l'}^{k,n}$.

2.2 The upper-layer problem

Now suppose we are at an upper-layer decision time τ. Our goal in the upper-layer problem is to determine the repositioning of mobile chargers among zones to achieve the maximum expected reward in the remaining operational horizon in the day. For this purpose, we propose a stochastic dynamic decision-making formulation for tackle the problem; this formulation contains the following components:

- *States*: At the decision time τ, the current system states include: the number of mobile chargers in each zone $z \in \mathcal{Z}$, denoted by V_z^τ; the number of unserved orders in each zone $z \in \mathcal{Z}$, denoted by U_z^τ. The state variable is aggregated as $\mathbf{S}^\tau \triangleq (\mathbf{V}^\tau, \mathbf{U}^\tau)$.
- *Actions*: At the decision time τ, the action is denoted by $A_{z,z'}^\tau$, indicating the number of mobile chargers repositioned from zone $z \in \mathcal{Z}$ to zone $z' \in \mathcal{Z}$. The aggregated form is denoted by $\mathbf{A}^\tau$.

When the state-action pair $(\mathbf{S}^\tau, \mathbf{A}^\tau)$ is given, the state of the next stage, i.e., $\mathbf{S}^{\tau+1}$, can be inferred accordingly. However, due to the intrinsic stochasticity of the operation process, $\mathbf{S}^{\tau+1}$ cannot be determined uniquely, so it is essentially a random variable. Compactly, we use ζ^τ to denote the random seed, and the combination $(\mathbf{S}^\tau, \mathbf{A}^\tau, \zeta^\tau)$ could uniquely identify the state of the next stage.

Let us denote the optimal expected reward associated with state $\mathbf{S}^\tau$ as $R^{\tau,*}(\mathbf{S}^\tau)$, and the current-stage reward under $(\mathbf{S}^\tau, \mathbf{A}^\tau, \zeta^\tau)$ is denoted by $Q^\tau(\mathbf{S}^\tau, \mathbf{A}^\tau, \zeta^\tau)$. Then, we can write down the optimality condition of the stochastic dynamic decision-making problem as:

$$R^{\tau,*}(\mathbf{S}^\tau) = \max_{\mathbf{A}^\tau} \ \mathbb{E}_{\zeta^\tau}\left[Q^\tau(\mathbf{S}^\tau, \mathbf{A}^\tau, \zeta^\tau) + R^{\tau+1,*}(\mathbf{S}^\tau, \mathbf{A}^\tau, \zeta^\tau)\right] \qquad \forall \tau \in \mathcal{T} \tag{13}$$

where $\mathcal{T}$ is the set of all upper-layer decision-making stages. Meanwhile, if we denote the optimal action given the state information at stage τ as $\mathbf{A}^{\tau,*}(\mathbf{S}^\tau)$, then we can write down the decision-making problem as below:

$$\mathbf{A}^{\tau,*}(\mathbf{S}^\tau) \in \arg\max_{\mathbf{A}^\tau} \ \mathbb{E}_{\zeta^\tau}\left[Q^\tau(\mathbf{S}^\tau, \mathbf{A}^\tau, \zeta^\tau) + R^{\tau+1,*}(\mathbf{S}^\tau, \mathbf{A}^\tau, \zeta^\tau)\right] \qquad \forall \tau \in \mathcal{T} \tag{14}$$

When the relocation decision is made, the execution of the actions can be with a simple manner. For instance, if the decision requires a zone to relocate 4 mobile chargers to other zones, then we can choose four chargers with the earliest available times to leave the zone immediately after they finish their current services. The destination nodes of these vehicles can be set as the ones with the smallest travel times in the trageted zones.

3 Results

To verify the effectiveness of our proposed framework, we conduct numerical experiments based on two networks, a two-zone toy network and a realistic network in Chongqing.

References

Ferrero, Enrico, Alessandrini, Stefano, & Balanzino, Alessia. 2016. Impact of the electric vehicles on the air pollution from a highway. *Applied energy*, **169**, 450–459.