Efficient Deployment of Electric Vehicle Charging Infrastructure: Simultaneous Optimization of Charging Station Placement and Charging Pile Assignment

Ying Zhang, *Member, IEEE*, Yanhao Wang, Fanyu Li, Bin Wu, Yao-Yi Chiang, and Xin Zhang, *Member, IEEE*

***Abstract*— Charging infrastructure deployment is to seek the proper plan of settling charging stations and charging piles under multiple constraints, such as recharging demand, cruising range, etc., and it has been asserted as an NP-Complete problem. In this paper, we propose a multicriteria-oriented approach of efficiently deploying charging infrastructure to cope with the problem. We firstly formulate five realistic charging objective functions that exhibit a significant diminishing returns effect, i.e., submodularity, and then exploit the submodularity of these objectives to design the acceleration algorithms for Charging Station PLacement (CSPL) with the provable performance guarantees. The corresponding algorithms are respectively named Lazy Greedy with Direct Gain (LGDG) and Lazy Greedy with Effective Gain (LGEG), and they scale well to the road networks of arbitrary size. Relying on the inference that the linear combination of submodular functions is still a submodular function, we treat CSPL as a multicriteria optimization problem that can be efficiently solved by the proposed algorithms. Moreover, we employ Erlang-Loss system to gain an optimal Charging Pile ASsignment (CPAS), which is capable of reducing the gap between the growing complexity of charging demands and the constrained supply of charging resources in considering the correlation between the primary human activities and the charging process. The experimental evaluation with real data sets shows that, compared with the state-of-the-art methods, the proposed approach reveals better effectiveness and efficiency, and it offers a potent solution to the planning of charging infrastructure for electric vehicles with large-scale datasets in reality.**

***Index Terms*— Submodularity, charging station placement, charging pile assignment, electric vehicle.**

I. Introduction

EMISSIONS from road transportation are responsible for most of air quality problems in highly populated urban areas. Replacing conventional Internal Combustion Engine (ICE) vehicles with Electric Vehicles (EVs) offers an appealing chance to reduce excessive energy dependence and to mitigate greenhouse gas emissions. However, "range anxiety" is a significant obstacle to limit drivers' acceptance for EVs, and it always arises the drivers' worry about being stranded with insufficient range to reach the destination. One feasible solution to alleviate such range anxiety is increasing public charging infrastructure, which is not well developed yet in many areas of the world.

To efficiently deploy the charging infrastructure, we propose a multi-criterion-oriented optimization approach integrating the solutions to placing charging stations and assigning charging piles. The essential of Charging Station PLacement (CSPL) is to find k constrained locations with the maximum gain, which is NP-hard and thus brings two key issues, i.e., the efficiency and the target problem granularity. The typical methodologies to solve CSPL include Mixed Integer Linear Programming (MILP), meta-heuristic approach, and greedy-based algorithms. We reviewed and analyzed the representative works of adopting the methodologies, based on which the comparison among the above three types of methodologies are summarized and listed according to the time cost of solving problem, i.e., the efficiency, and the treatable problem size, i.e., the target problem granularity. We also categorized the methodologies into deterministic and probabilistic. That a methodology is deterministic means the results would be identical through multiple runs of the same problem instance, whereas a methodology of probabilistic nature would produce alterable outcomes during multiple runs. Elicited from Table I, we designed two accelerating algorithms (named LGDG and LGEG) based on the greedy methods with incorporating the submodularity [12], [13] of the objectives for CSPL.

TABLE I

Comparison Among the Typical Methods

Method	Efficiency	Treatable size	Category	Representative works
MILP	Low	Low	Deterministic	[1], [2]
Meta-heuristic	Meduim	Medium	Probabilistic	[3], [4], [5], [6], [7]
Greedy	High	High	Deterministic	[8], [9], [10]

Referring to assigning charging piles given the placed charging stations, we employ Erlang-Loss system to gain an optimal Charging Pile ASsignment (CPAS), which is capable of reducing the gap between the growing complexity of charging demands and the constrained supply of charging resources in considering the correlation between the primary human activities and the charging process.

II. The Solutions to CSPL Problem

A. Problem Statement

Problem Definition: Given a road network $G =< V, E >$, a set of objective functions $\{F_1, \ldots, F_L\}$, a budget of k charging stations with N new charging piles in total, where value $k \leq |V|$, we aim to find the optimal k-location set A to build charging stations and the optimal assignment of charging piles to the stations, so as to maximize the total gain under multiple criteria. In other words, we need to pick proper weights $\lambda_1 > 0$, $\lambda_2 > 0$,..., $\lambda_L > 0$, to optimize the total benefit from $F = \sum_{i=1}^{L} \lambda_i \times F_i(A)$, where $\sum_{i=1}^{L} \lambda_i = 1$.

B. The Objectives to CSPL Problem

Objective 1 Maximizing the Charging Likelihood (MCL): Parking-lot-based charging station placement is more proper for the long-duration battery charging because the charging process is just one

Manuscript received April 11, 2019; revised September 11, 2019, December 13, 2019, April 11, 2020, and April 15, 2020; accepted April 21, 2020. This work was supported in part by the Fundamental Research Funds for the Central Universities under Grant 2018MS024, in part by the National Natural Science Foundation of China under Grant 61305056, and in part by the Jilin Provincial Science and Technology Planning Project under Grant 20190303133SF. The Associate Editor for this article was Y. Wang. *(Corresponding author: Ying Zhang.)*

Ying Zhang, Yanhao Wang, Fanyu Li, and Bin Wu are with the School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China (e-mail: ying.zhang@ieee.org).

Yao-Yi Chiang is with the Spatial Sciences Institute, Dana and David Dornsife College of Letters, Arts and Sciences, University of Southern California, Los Angeles, CA 90007-0374 USA.

Xin Zhang is with the School of Computer Science and Technology, Changchun University of Science and Technology, Changchun 130022, China.

Digital Object Identifier 10.1109/TITS.2020.2990694

of the incidental activities during a trip (e.g., sightseeing, shopping, etc.) and does not require extra waiting time. The parking lots connecting with a large number of tracks should thus be selected. The corresponding criterion is defined as the following expression:

$$\begin{aligned} Z &= \max(R(A) : A \subseteq V, |A| = k) \\ R(A) &= \sum\nolimits_{t \in T} w(t) \times R(A,t) \\ R(A,t) &= \begin{cases} 1, & \exists v \in A, passby(v,t) = true \\ 0, & otherwise \end{cases} \end{aligned} \tag{1}$$

In Formula (1), $R(\bullet)$ denotes the charging likelihood function which is the same in all the presented objective functions, A is a subset of V in the given road network $G =< V, E >$, k denotes the maximum number of charging stations selected according to the total financial budget and T is a set of trajectories. $w(t)$ represents the weight (importance) of trajectory t, and it is set to 1.0 in this work. $passby(v,t)$ is used to determine whether $v \in t$. From Formula (1), the higher the value of $R(A)$ is, the more probable that the set of parking-lots A should be chosen.

Objective 2 Maximizing the Charging Willingness of Drivers (MCW): To save driving time during a trip, drivers prefer to charge the vehicles nearby either the start point or the destination point rather than in the middle of the trip. We suppose that drivers' willingness obeys a normal distribution described by Formula (2). For the normal distribution, the probability that a normal deviate lies in the range between $\mu - 3\sigma$ and $\mu + 3\sigma$ is given by $P\{|X - \mu| < 3\sigma\} = 0.9974$. It means the probability that X falls outside $(\mu - 3\sigma, \mu + 3\sigma)$ is less than three thousandths, i.e., the "3σ" principle of the normal distribution. In our work, "nearby" refers to the area of $1km^2$ surrounding the given point, and 3σ should be correspondingly set as approximately 1.0. As shown in Formula (2), we set $\sigma = \frac{1}{\sqrt{2\pi}} \approx 0.4$ because it happens to be a type of standard normal distribution with variance $\sigma^2 = \frac{1}{2\pi}$.

$$f(x) = \begin{cases} \frac{1}{\sigma \cdot \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2 \cdot \delta^2}}, \mu = 0, \sigma = \frac{1}{\sqrt{2\pi}}, & 0 \le x \le \frac{L}{2} \\ \frac{1}{\sigma \cdot \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2 \cdot \delta^2}}, \mu = L, \sigma = \frac{1}{\sqrt{2\pi}}, & \frac{L}{2} \le x \le L \end{cases} \tag{2}$$

Given the probability density function in Formula (2), we define the second objective by Formula (3). $dis(v,t)$ represents the distance from the start point of a trajectory t to a specific location v. Function $f(\bullet)$ is given by Formula (2).

$$\begin{aligned} Z &= \max(R(A) : A \subseteq V, |A| = k) \\ R(A) &= \sum\nolimits_{t \in T} w(t) \times R(A,t) \\ R(A,t) &= \max\{f(dis(v,t)) : v \in A\} \end{aligned} \tag{3}$$

Objective 3 Maximizing the Charging Demand of EVs : For EVs, the charging demand grows with the distance from the start point increasing. The following criterion is used to select the locations for charging stations. $length(t)$ denotes the length of a trajectory t.

$$\begin{aligned} Z &= \max(R(A) : A \subseteq V, |A| = k) \\ R(A) &= \sum\nolimits_{t \in T} w(t) \times R(A,t) \\ R(A,t) &= dis\_ratio(A,t)) \\ & dis\_ratio(A,t) \\ &= \max\{dis(v,t)/length(t) : v \in A\} \end{aligned} \tag{4}$$

Objective 4 Maximizing the Coverage of POIs (MCP): As the charging process is an auxiliary process of the main activities occurring at POIs, the density of POIs around the parking lots is meaningful for drivers/travelers to choose their parking locations. Formula (5) specifies the corresponding objective function, in which POI is a set of vertices representing POIs, $w(i)$ denotes the weight of vertex i, and *threshold* is designated to prescribe the near area.

$$\begin{aligned} Z &= \max(R(A) : A \subseteq V, |A| = k) \\ R(A) &= \sum\nolimits_{i \in POI} w(i) \times R(A,i) \\ R(A,i) &= \begin{cases} 1, & \exists v \in A, dis(v,i) \le threshold \\ 0, & otherwise \end{cases} \end{aligned} \tag{5}$$

Objective 5 Maximizing the Distance Reduction from CS to POIs (MDR): Since the principal activities usually occur at POIs during a trip, minimizing the distance from CS to the near POIs is necessary and meaningful for drivers to save the time cost in dealing with charging. The corresponding objective function is specified as Formula (6), in which $threshold$ has the same meaning as that in Formula (5).

$$\begin{aligned} Z &= \max(R(A) : \text{A} \subseteq V, |A| = k) \\ R(A) &= \sum\nolimits_{i \in POI} w(i) \times R(A,i) \\ R(A,i) &= threshold - \tau(A,i) \\ \tau(A,i) &= \min\{\min_{v \in A} dis(v,i), threshold\} \end{aligned} \tag{6}$$

C. Problem Analysis

All the presented objective functions share several intuitive properties. Given that $R(A)$ is non-negative for each placement A (seen in Formula (1, 3, 4, 5, 6), we generally want to maximize the benefit from each distinct function. If we place no CS, the total benefit we can get is 0, i.e., $R(\emptyset) = 0$. We can also find that each function R is non-decreasing, i.e., for subsets $A \subseteq B \subseteq S$, it holds that $R(A) \le R(B)$, hence the benefit can only increase if we place more charging stations. Besides, there is an additional intuitive property: when we add a CS to a large-scale set of the deployed charging stations, we would get less benefit than that gained by adding one CS to a small-scale set. Such diminishing returns are formalized as *submodularity* by G.L. Nemhauser *et al.* in [12], i.e., a set function F is called submodular if for all subsets $A \subseteq B \subseteq S$ and the element $s \in S$ it holds that $F(A \cup \{s\}) - F(A) \ge F(B \cup \{s\}) - F(B)$. Referring to the proof of submodular in [12], we can prove that all the objective functions mentioned above are submodular, and the linear combination of submodular functions is still a submodular function. In the following sections, we will employ the submodularity of the presented objective functions accelerate solving the corresponding optimization problem.

D. The Acceleration Algorithms for CSPL Problem

Simple Greedy algorithm is a conventional heuristic method and provides the provable guarantees that the greedy solutions achieve at least $(1 - 1/e) \approx 63\%$ times of the optimal results. Minoux in [8] and Jure Leskovec *et al.* in [10] respectively introduced the Lazy Greedy (LG) algorithm and the CELF (Cost-Effective Lazy Forward selection) algorithm to separately improve the efficiency of Simple Greedy algorithm. To handle the issue that the computing efficiency of the algorithms would be suppressed under large-scale datasets, we propose two improved algorithms (respectively abbreivated as LGDG and LGEG) that are independent with the dataset scale to accelerate both the Lazy Greedy algorithm and the CELF algorithm when dealing with the CSPL problem under large-scale datasets.

As stated in Algorithm 1, LGDG divides the trajectory dataset (denoted as T) into two subsets, i.e., the common trajectories covered by both the placement A and the vertex v^*, and the trajectories

Algorithm 1 Lazy Greedy With Direct Gain: LGDG

```
A ← ∅;
//AT: a set of trajectories covered by A;
//Tv*: a set of trajectories covered by v*;
for each v ∈ V do
    δ_v ← ∞;
while ∃v ∈ V\A and c(A ∪ {v}) ≤ Budget do
    for each v ∈ V\A do
        cur_v ← false;
    while true do
        v* ← argmax_{v∈V\A, c(A∪{v})≤Threshold} δ_v;
        if cur_{v*} then
            A ← A ∪ v*;
            for each t ∈ Tv* do
                R(A, t) = max{R(A − v*, t), R({v*}, t)};
            break;
        else
            δ_{v*} ← Σ_{t∈(AT∪Tv*)} [P(t) × (R(A ∪ {v*}, t) − R(A, t))]
                   + Σ_{t∈(Tv*−(AT∪Tv*))} [P(t) × R(A ∪ {v*}, t)];
            cur_{v*} ← true;
return A;
```

Algorithm 2 Lazy Greedy With Effective Gain: LGEG

```
A ← ∅;
for each v ∈ V do
    δ_v ← +∞;
while ∃v ∈ V\A and c(A ∪ {v}) ≤ Budget do
    for each v ∈ V\A do
        cur_v ← false;
    while true do
        v* ← argmax_{v∈V\A, c(A∪{v})≤Threshold} δ_v;
        if cur_{v*} then
            A ← A ∪ v*;
            for each t ∈ Tv* do
                R(A, t) = max{R(A − v*, t), R(v*, t)};
            break;
        else
            δ_{v*} ← Σ_{t∈(AT∩Tv*), R({v*},t)>R(A,t)} [P(t) × (R({v*}, t) − R(A, t))]
                   + Σ_{t∈(Tv*−(AT∩Tv*))} [P(t) × R({v*}, t)];
            cur_{v*} ← true;
return A;
```

belonging to Tv^* but not AT, in which AT represents the set of trajectories covered by the placement A and Tv^* denotes the set of trajectories covered by the vertex v^*. Correspondingly, we can directly get the benefits from adding one charging station v^* according to the equation $\delta_{v^*} = \sum_{t\in(AT\cap Tv^*)} P(t) \times (R(A \cup \{v^*\}, t) - R(A, t)) + \sum_{t\in(Tv^*-(AT\cap Tv^*))} P(t) \times R(A \cup \{v^*\}, t)$ without simultaneously calculating both $R(A \cup \{v^*\})$ and $R(A)$. Since the set Tv^* only accounts for a small part of the entire trajectory collection (i.e., T), Algorithm 1 enables to improve the efficiency of deploying charging stations dramatically.

According to the submodularity revealed by the objective functions, Algorithm 2 divides the trajectory dataset into two subsets, and narrows the search scope within the first set by eliminating those trajectories with which the benefit gained by the placement A is less than that of the vertex v^*. Namely, if $t \in AT \cap Tv^*$ and $R(A, t) > R(v^*, t)$, then we could correspondingly remove t from the first group and assess the marginal gain with $\delta_{v^*} = \sum_{t\in AT\cap Tv^*, R(\{v^*\},t)>R(A,t)} P(t) \times (R(\{v^*\}, t) - R(A, t)) + \sum_{t\in(Tv^*-(AT\cap Tv^*))} P(t) \times R(\{v^*\}, t)$ directly. The evaluation of the algorithms is depicted in Section IV.

E. Multicriteria Optimization

Different objective functions produce a variety of deployments, and each deployment A has a score vector denoted as $R(A) = (R_1(A), \ldots, R_m(A))$, where m is the number of objectives. In practical situations, two deployments A_i and A_j might be incomparable, i.e., given $R_i(A_1) > R_i(A_2)$, there could still exist a fact that $R_j(A_1) < R_j(A_2)$. We prospect to simultaneously optimize multiple objectives, namely, to achieve the trade-off among all the given objectives. In this paper, we calculate the Pareto optimal solutions for trading off the objective functions because it is generally impossible to seek out the globally optimal solution to a multi-criterion optimization problem. The non-negative linear combination of objective functions can transform one multi-criterion optimization problem into a scalarized multi-objective optimization problem as depicted by Formula (7) [15]. Correspondingly, a Pareto frontier [16] consisting of a set of Pareto optimal solutions can be obtained by the ad hoc adjustment of the weights $\{\lambda_i\}$ in Formula (7). Although the number of generated Pareto solutions grows to be huge in practice, LGDG and LGEG can help improve the efficiency of obtaining Pareto solutions since the combined function $R(A)$ is a linear combination of submodular.

$$\max R(A) \;\; R(A) = \sum_i \lambda_i R_i(A) \quad (\lambda_i > 0) \tag{7}$$

III. The Solutions to CPAS Problem

Public charging is usually an affiliated activity during a trip so that the actual charging time would mainly depend on the primary activities of the trip. Seen that the remaining time of acting the activities is unpredictable, we assume that drivers would not stop to queue at CS when none of the charging piles are available, by which the latest arrived EV would lose the chance of charging if the charging piles are all occupied [14]. We employ the M/M/n/n Queuing System to solve the Optimal Charging Pile Assignment. The most important evaluation index of the queuing system is related to the proportion of losing EVs at each CS. We set P_{loss} as the measurement. Given a predefined threshold P^*_{loss} for the loss rate at charging stations, the optimal charging pile assignment can be obtained through the following objective function:

$$\begin{aligned} \min \;&: n, \quad n \in N \\ s.t. \;&: P_{loss} \le P^*_{loss} \end{aligned} \tag{8}$$

i.e.,

$$\begin{aligned} &\min : n, \quad n \in N \\ &s.t. \;: P_{loss} = P_n = \frac{\frac{\rho^n}{n!}}{\sum_{k=0}^{n} \frac{\rho^k}{k!}} \le P^*_{loss} \end{aligned} \tag{9}$$

where P_n represents the steady state probability when there are n EVs at the charging station. By the regular experience, the phenomena of EVs arriving at CS obeys a Poisson process with the rate λ, and the random charging time is negative exponentially distributed with the rate μ. $\rho = \frac{\lambda}{\mu}$ denotes the traffic intensity or offered load. In addition to the M/M/n/n Queue, we also adopt other three methods to assign charging piles in Section IV.

IV. Experimental Evaluation

Since the personal privacy protection brings collecting the private vehicle driving data for analyzing and assessing public charging demands with numerous constraints and barriers, we choose to utilize the public trajectory data collected from the taxi cabs in Beijing for building up and verifying our research work. Such alteration is made based on the hypothesis that the travelers'/drivers' behaviors remain unchanged when switching to drive/take electric vehicles.

A. Dataset and Settings

1) Real Datasets: Road Networks: To assess the efficiency and scalability of the proposed methods, we extract the road network from Beijing. The bounding box is [(116.092, 40.117), (116.707, 39.685)], which covers an area of 2, 520 km^2 and contains 83, 917 vertices as well as 110, 114 road segments.

Taxi Trajectories: The GPS trajectory dataset is generated by collecting the driving data from 33, 000 taxicabs from Beijing during a 87-day period. The dataset contains 268, 791 trajectories (segmented corresponding to the taxi orders).

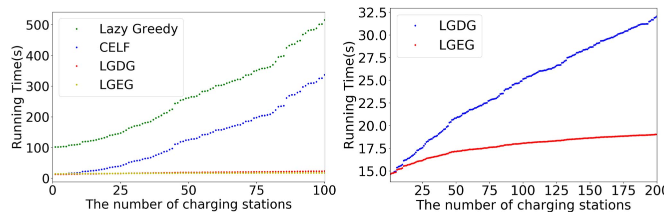
Parking Lot Dataset: We extract the dataset of the public parking lots from the government website with the information about name, type, and address. There are 2, 087 public parking lots within the selected demonstrative area.

2) Metrics: We suggest two metrics to evaluate the proposed solutions.

- *Coverage rate* refers to the proportion of EVs that can get charged to the total charging demands.
- *Utilization rate* represents the ratio of the time each charging pile used per day to the total working time of the day.

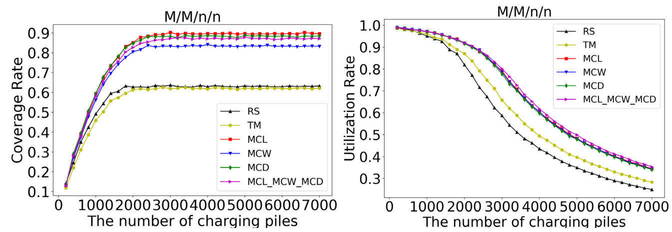
B. Evaluation of Charging Station Placement

1) Efficiency Comparison: To evaluate the efficiency of the proposed algorithms with a large-scale dataset, we compared our work with two baselines within the selected demonstrative area. As shown in Figure 1(a), CELF, LGDG, and LGEG are all better than the conventional Lazy Greedy Algorithm. With the number of charging stations increasing gradually, both LGDG and LGEG outperform CELF significantly. With the amount of charging stations increasing, the running time of CELF rises a lot, whereas the time cost of our proposed algorithms keeps steady. Because CELF Algorithm traverses all the trajectories covered by the placement A and the station $v*$, the number of trajectories needing traversing would increase as the size of placement A enlarges. In contrast, our algorithms only traverse the trajectories passing by the newly added station $v*$ and thus get little impact when the size of placement A increases. Moreover, as shown in Figure 1(b), LGEG gets better performance than LGDG as it further narrows the traversing scope by eliminating those trajectories on which the obtained benefit gain from placement A is less than that from the newly added station $v*$.



(a) Efficiency comparisons among different charging station deployments (b) Efficiency comparisons between LGDG and LGEG

Fig. 1. Efficiency comparisons.



(a) Comparison of the coverage rate among various charging station deployments (b) Comparison of the utilization rate among various charging station deployments

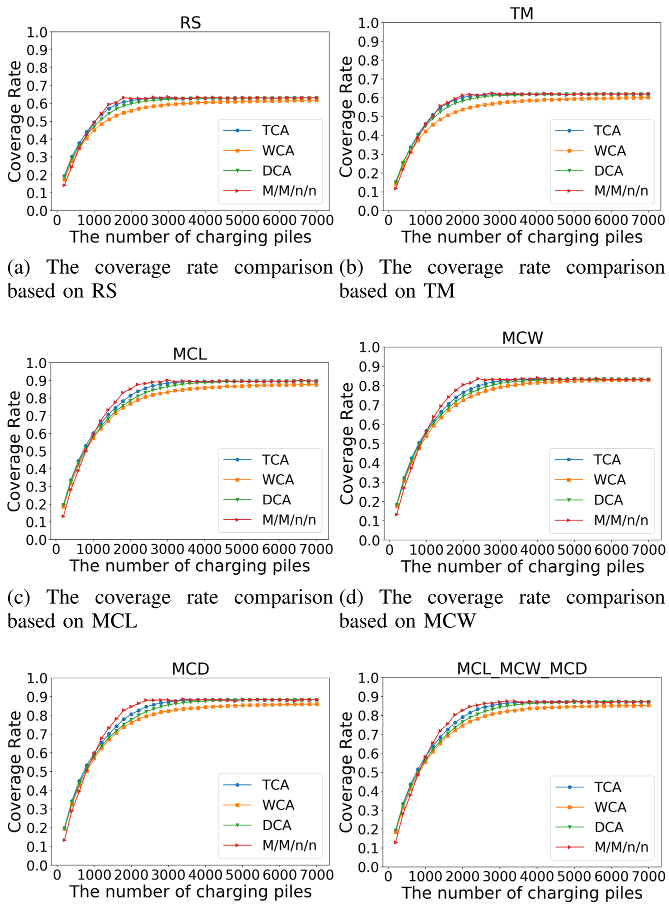
Fig. 2. Performance comparison for CSPL.

2) Performance Comparison for CSPL: In this section, we evaluate the performance of our proposed algorithms of placing charging stations in terms of the coverage rate and the utilization rate. We assume that each trajectory is formed by one EV, and the average speed of EV is $50km/h$. Then, we can calculate the time cost of each EV to reach charging stations along the trajectories. Besides, λ denoting the mean arrival rate of EVs into the queuing system can be calculated. For all the EVs within one CS, we assume their charging time is negative exponentially distributed with the rate μ according to the POI types that the current CS is close to. Table II lists the average staying time of EVs at each type of POIs where charging stations are located. We take the queuing theory model to calculate the number of charging piles for each CS with the above mentioned λ and μ and list the coverage rate and utilization rate of each method for CSPL.

We compare six methods for CSPL problem with $|A| = 200$ (seen in Figure 2. In the figure, *RS* represents the way to randomly select the locations for charging stations, and *TM* refers to picking the first charging stations through sorting the number of the covered trajectories in descending order. Given MCL, MCW, and MCD discussed above, MCL_MCW_MCD represents the multicriteria optimization combining the three individual objectives. As shown in Figure 2, our proposed methods outperforms RS and TM in terms of both the coverage rate and the utilization rate. With the coverage rate as the evaluation criterion, MCL could easily get a great reputation because it tends to select the locations with relatively large coverage, and it thus gets the best performance as illustrated in Figure 2(a). Meanwhile, we also incorporate the utilization rate as a second metric to highlight the improvement in the work. Through analyzing the experimental results, we find that the method for multicriteria optimization, i.e., MCL_MCW_MCD, plays the best performance (seen in Figure 2(b)) because it synthesizes the solutions to MCL, MCW, and MCD at the same time and suggests a trade-off among the multiple objectives.

TABLE II
THE DISTRIBUTION OF THE STAY TIME IN POIS

POI types	*Store*	*Restaurant*	*Scenic Spot*	*Vehicle Service*	*Entertainment*	*Others*
$Average\ stay\ time\ (Unit:\ hour)$ $t = \frac{1}{u}$	2.5	1	4	4	3	1

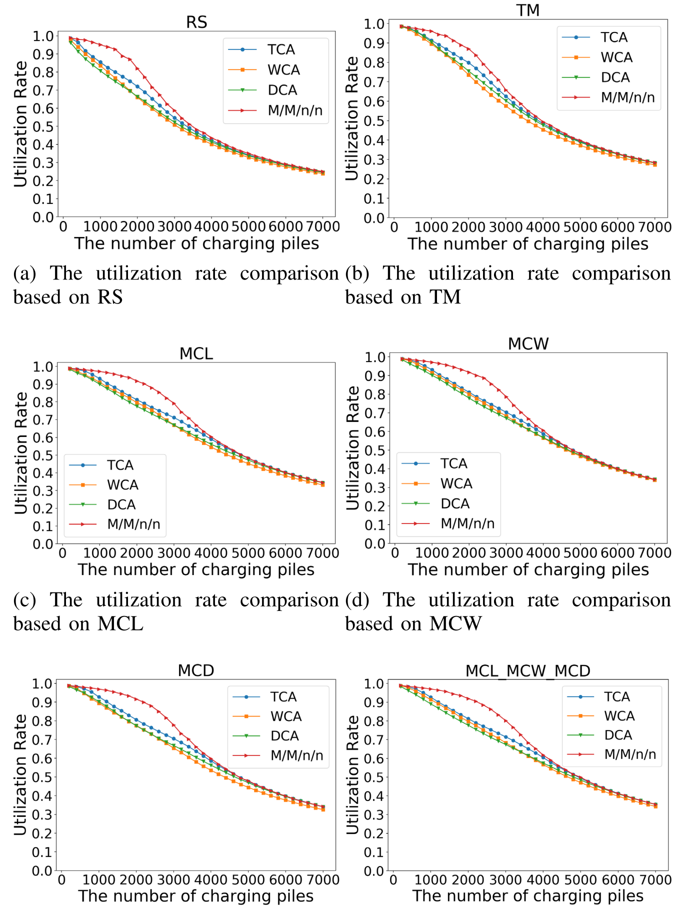


(a) The coverage rate comparison based on RS (b) The coverage rate comparison based on TM

(c) The coverage rate comparison based on MCL (d) The coverage rate comparison based on MCW

(e) The coverage rate comparison based on MCD (f) The coverage rate comparison based on MCL_MCW_MCD

Fig. 3. The comparisons of coverage rate among various strategies of charging pile assignment.



(a) The utilization rate comparison based on RS (b) The utilization rate comparison based on TM

(c) The utilization rate comparison based on MCL (d) The utilization rate comparison based on MCW

(e) The utilization rate comparison based on MCD (f) The utilization rate comparison based on MCL_MCW_MCD

Fig. 4. The comparisons of utilization rate among various strategies of charging pile assignment.

C. Evaluation of Charging Pile Assignment

In this section, we evaluate four methods of assigning charging piles. TCA (*Trajectory-based Charging pile Assignment*) is the assignment strategy based on the coverage ratio. The number of charging piles assigned to CS_i is according to $\frac{Coverage(V_i)}{Coverage(A)} \times N$, where $Coverage(V_i)$ designates the number of trajectories passing by CS_i, and $Coverage(A)$ represents the number of trajectories covered by placement A, and N denotes the total number of charging piles. DCA (*charging Demands-based Charging pile Assignment*) represents the assignment strategy based on the coverage ratio of charging demands $\frac{MCD(V_i)}{MCD(A)}$, and WCA (*charging Willingness-based Charging pile Assignment*) represents the assignment strategy based on the coverage ratio of charging willingness $\frac{MCW(V_i)}{MCW(A)}$. M/M/n/n uses the queuing theory model to get the number of charging piles for each CS.

Figure 3 states the coverage rate comparisons among TCA, DCA, WCA, and M/M/n/n under the preset CSPL strategies. Figure 4 shows the comparisons among the four methods of CPAS in terms of the utilization rate. We find that the proposed method, i.e., M/M/n/n, outperforms the other methods according to both Figure 3 and Figure 4. Particularly, when the total number of charging piles is less than 4000, our proposed method is much better than the others. In reality, the initial number of charging piles should usually be small due to the constrained budget. So it is usually recommended that each CS is configured with ten charging piles. In such a situation, M/M/n/n could improve the utilization rate by 10% over the other methods, which proves that the proposed method has good prospects in practice.

V. CONCLUSION

In this paper, we proposed a novel approach to efficiently deploy electric vehicle charging infrastructure toward improving the convenience and extending the cruising range of electric vehicles. Our main contributions include: (1) the formulation of the five realistic charging objective functions exhibiting submodularity; (2) the effective heuristic algorithms, i.e., LGDG and LGEG, to accelerate the conventional algorithms for optimizing submodular functions under large scale dataset; (3) the efficient solution to the multicriteria optimization problem of CSPL; (4) the methods of assigning charging piles for improving the piles' service efficiency and feeding heavier charging load during the service period. The experimental results obtained with acting the research work with the real large-scale datasets in Beijing indicates that our approach can provide an effective and efficient deployment of electric vehicle charging infrastructure.

References

[1] J. Yang, J. Dong, and L. Hu, "A data-driven optimization-based approach for siting and sizing of electric taxi charging stations," *Transp. Res. C, Emerg. Technol.*, vol. 77, pp. 462–477, Apr. 2017.

[2] A. Y. S. Lam, Y.-W. Leung, and X. Chu, "Electric vehicle charging station placement," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Oct. 2013, pp. 510–515.

[3] S. Mehar and S. M. Senouci, "An optimization location scheme for electric charging stations," in *Proc. Int. Conf. Smart Commun. Netw. Technol. (SaCoNeT)*, Jun. 2013, pp. 1–5.

[4] P.-S. You and Y.-C. Hsieh, "A hybrid heuristic approach to the problem of the location of vehicle charging stations," *Comput. Ind. Eng.*, vol. 70, pp. 195–204, Apr. 2014.

[5] P. Phonrattanasak and N. Leeprechanon, "Optimal placement of EV fast charging stations considering the impact on electrical distribution and traffic condition," in *Proc. Int. Conf. Utility Exhib. Green Energy Sustain. Develop. (ICUE)*, Mar. 2014, pp. 1–6.

[6] G. Gatica, G. Ahumada, J. W. Escobar, and R. Linfati, "Efficient heuristic algorithms for location of charging stations in electric vehicle routing problems," *Stud. Informat. Control*, vol. 27, no. 1, pp. 73–82, Mar. 2018.

[7] W. Dait, Y. Lit, X. Gan, and G. Xie, "Fast charging station placement with elastic demand," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–7.

[8] M. Minoux, "Accelerated greedy algorithms for maximizing submodular set functions," in *Optimization Techniques*. Berlin, Germany: Springer, 1978, pp. 234–243.

[9] X. Wang, C. Yuen, N. U. Hassan, N. An, and W. Wu, "Electric vehicle charging station placement for urban public bus systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 1, pp. 128–139, Jan. 2017.

[10] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen, and N. Glance, "Cost-effective outbreak detection in networks," in *Proc. 13th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2007, pp. 420–429.

[11] Y. Xiong, J. Gan, B. An, C. Miao, and A. L. C. Bazzan, "Optimal electric vehicle fast charging station placement based on game theoretical framework," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 8, pp. 2493–2504, Aug. 2018.

[12] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, "An analysis of approximations for maximizing submodular set functions—I," *Math. Program.*, vol. 14, no. 1, pp. 265–294, Dec. 1978.

[13] G. Nemhauser and L. Wolsey, *Studies on Graphs and Discrete Programming*. Amsterdam, The Netherlands: North Holland, , 1981, pp. 279–301.

[14] J. Sztrik, *Basic Queuing Theory*. Debrecen, Hungary: Univ. Debrecen, 2012.

[15] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[16] A. V. Lotov and K. Miettinen, "Visualizing the Pareto frontier," in *Multiobjective Optimization* Berlin, Germany: Springer, 2008, pp. 213–243.