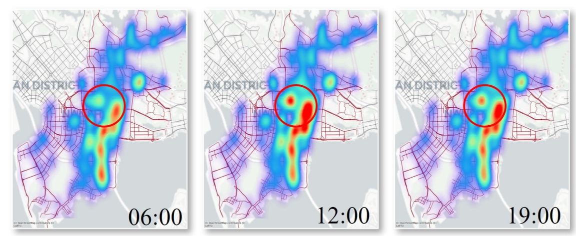
Reinforcement Learning for Hybrid Charging Stations Planning and Operation Considering Fixed and Mobile Chargers

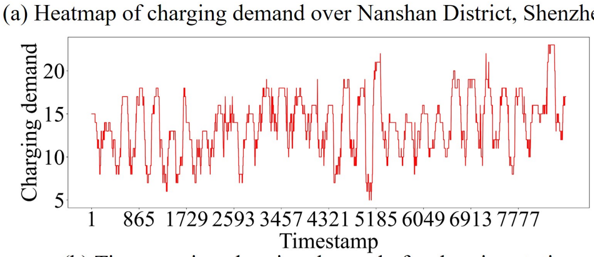
Yanchen Zhu[1] , **Honghui Zou**[2] , **Chufan Liu**[1] , **Yuyu Luo**[1] , **Yuankai Wu**[3] and **Yuxuan Liang**[1,*]
[1]Hong Kong University of Science and Technology (Guangzhou)
[2]Beihang University
[3]Sichuan University
yzhu156@connect.hkust-gz.edu.cn, hhzou@buaa.edu.cn, cliu779@connect.hkust-gz.edu.cn,
yuyuluo@hkust-gz.edu.cn, Kaimaogege@gmail.com, yuxliang@outlook.com

Abstract

The success of vehicle electrification relies on efficient and adaptable charging infrastructure. Fixed-location charging stations often suffer from underutilization or congestion due to fluctuating demand, while mobile chargers offer flexibility by relocating as needed. This paper studies the optimal planning and operation of hybrid charging infrastructures that combine both fixed and mobile chargers within urban road networks. We formulate the Hybrid Charging Station Planning and Operation (HCSPO) problem, jointly optimizing the placement of fixed stations and the scheduling of mobile chargers. A charging demand prediction model based on Model Predictive Control (MPC) supports dynamic decision-making. To solve the HCSPO problem, we propose a deep reinforcement learning approach enhanced with heuristic scheduling. Experiments on real-world urban scenarios show that our method improves infrastructure availability—achieving up to 244.4% increase in coverage—and reduces user inconvenience with up to 79.8% shorter waiting times, compared to existing solutions.



(a) Heatmap of charging demand over Nanshan District, Shenzhen



(b) Time-varying charging demand of a charging station

Figure 1: Dynamic attribute of charging demand

1 Introduction

The transition towards vehicle electrification is rapidly advancing, driven by its substantial societal and environmental benefits. However, the widespread adoption of electric vehicles (EVs) hinges on the availability and efficiency of charging infrastructure [von Wahl *et al.*, 2022], which requires a comprehensive and adaptable planning approach to accommodate the increasing demand. A critical challenge in this process is accounting for the dynamic nature of charging demand, which is often overlooked in traditional station planning methods.

As shown in Figure 1, panel (a) presents a heat map of charging demand across Nanshan District in Shenzhen, China, highlighting how demand distribution varies throughout the day, from 6:00 AM to 7:00 PM. Meanwhile, panel (b) depicts the time series of charging demand for a specific station in Shenzhen, further illustrating the fluctuations in demand at a granular level. These spatio-temporal patterns collectively emphasize the dynamic nature of charging demand in urban areas.

While traditional charging stations with fixed-location chargers offer cost and grid stability benefits, their spatial rigidity leads to inefficiencies, especially in addressing demand-supply mismatches from fluctuating demand. For example, sizing infrastructure for peak demand causes underutilization during off-peak hours, while sizing for off-peak leads to congestion during peaks. These limitations highlight the need for adaptive infrastructure that can reconfigure in response to changing demand.

Emerging mobile chargers (MCs) have recently gained attention as a flexible supplement to urban charging networks [Afshar *et al.*, 2021]. Unlike fixed chargers, mobile chargers can be dynamically scheduled to align with changing demand patterns, offering a promising solution to the limitations of traditional charging infrastructure. In addition to addressing the limitations of fixed charging stations, the integration of MCs can also benefit the configuration of fixed charging stations. For example, in areas with demand surges, MCs can

*Corresponding author. Email: yuxliang@outlook.com

be scheduled for temporary service, eliminating the need for installing fixed stations in these locations. This approach ensures that fixed stations are deployed where demand is more consistent, improving overall system efficiency.

Considering optimization for both fixed and mobile charging infrastructures, this paper introduces a novel hybrid charging station planning and operation (HCSPO) problem, which aims to integrate the optimal planning of fixed charging stations with dynamic operation of mobile chargers. Specifically, we formulate the HCSPO problem within urban road networks to optimize the locations and configurations of fixed charging stations and to dynamically schedule mobile chargers to support real-time charging needs.

To enhance decision-making, we incorporate a charging demand prediction model based on Model Predictive Control (MPC) to provide foresight into demand fluctuations. Our solution leverages a reinforcement learning (RL) method, augmented with heuristic techniques, to adaptively optimize the planning and operation of the HCSPO problem. This integrated approach effectively bridges the strategic planning of fixed charging stations with the operational dynamics of mobile chargers, contributing to the sustainability and accessibility of urban transportation networks.

In summary, the contributions of this paper are as follows:

- We formulate the HCSPO problem into a rolling horizon framework within road networks, integrating the planning of fixed stations with mobile charging operations. Our approach incorporates a demand prediction model based on an MPC policy to enhance decision-making.
- We propose an adaptive RL algorithm, enhanced with heuristic scheduling techniques, to efficiently tackle the HCSPO problem, ensuring optimal planning and operation of both fixed and mobile infrastructures.
- We conduct extensive experiments on real-world dataset, which demonstrates our approach outperforms other baselines, offering superior societal benefits, including improved sustainability and user satisfaction.

2 Related Work

This section reviews related work on charging station planning and mobile charging infrastructure operations.

EV charging station planning is often modeled as a variant of the Facility Location Problem. However, many approaches rely on simplified assumptions about charging demand [Kchaou-Boujelben, 2021], which may not reflect actual spatio-temporal distributions. Some studies simulate EV recharging behavior using real-world datasets like GPS records [Cai *et al.*, 2014; Li *et al.*, 2017; Shahraki *et al.*, 2015; Yang *et al.*, 2017; Hwang *et al.*, 2015], while others, such as [Li *et al.*, 2017; Tu *et al.*, 2016], use taxi trajectories. Recent research also explores dynamic and stochastic environments [Wang *et al.*, 2023; Kchaou Boujelben and Gicquel, 2019; Kchaou-Boujelben and Gicquel, 2020; Lee and Han, 2017; Kong *et al.*, 2019; Yin *et al.*, 2023; Xiang *et al.*, 2019].

Mobile charging station placement is gaining traction. For example, [Tang *et al.*, 2020] proposes a two-phase framework to determine MCS placement, scheduling, and relocation, which is widely adopted in MCS placement and scheduling studies, including depot and fleet locations [Ting *et al.*, 2024; Liu *et al.*, 2024; Li *et al.*, 2024; Beyazıt and Taşcıkaraoğlu, 2023; von Wahl *et al.*, 2023]. Some studies validate approaches with real-world data, including GPS trajectories [Ting *et al.*, 2024; Liu *et al.*, 2024]. [Liu *et al.*, 2024] uses a Markov Decision Process (MDP) to reduce delays and increase the proportion of charged EVs, while [Ting *et al.*, 2024] extends the two-phase method with a multi-agent RL algorithm for dynamic MCS operation.

While substantial research exists on both fixed-location and mobile charging infrastructures, to our best knowledge, no study has integrated these within a single framework. This paper proposes an adaptive RL approach that simultaneously determines fixed station locations and configurations while dynamically scheduling mobile chargers.

3 Problem Statement

In this section, we formally define the HCSPO problem and provide the following key definitions:

Definition 1 (Road Network)**.** *Let $G = (V, E)$ be a directed weighted graph with V the set of vertices and E the set of edges. The vertices are the road network junctions, while the edges represent the roads direction-wise.*

Definition 2 (Dynamic Charging Demand)**.** *Considering the time-varied nature of charging demand, Let the recording $dem^{1,\dots,|T|}(v) = \left[dem^1(v), dem^2(v), \cdots, dem^{|T|}(v)\right]$ denote the charging demand of vertex v over the operation horizon divided into $|T|$ time slots.*

Definition 3 (Charging Station)**.** *A charging station (CS) s within the road network G is defined as a tuple $s = (v, x)$, where $v \in V$ is the location node, $x = (x_1, \dots, x_n)$ is a vector of length $n \in \mathbb{N}$ with $x_i \in \mathbb{N}$ being the number of fixed chargers of type i at s. We denote the set of all possible CSs as S. We set a limit on the number of fixed chargers of each s by $\sum_i x_i \leq K$.*

Definition 4 (Mobile Charger and Depot)**.** *A mobile charger (MC) m can be defined as $m = (l_m^{1,\cdots,|T|}, \tau_m^{1,\cdots,|T|}, q_m^{1,\cdots,|T|})$, which denotes the current location, accessible time (arrival time after scheduling decision) and remaining electricity of m at different time slots $t \in T$. Each schedule for MCs is batched with size of k_{MC} as a fleet. In addition, we define mobile depots $j \in J$ as the initial locations of MCs before scheduling and where they return to recharge when their batteries deplete, i.e., $l_m^1 = j \in J$.*

Definition 5 (Charging Plan)**.** *A charging plan $P = (S, M)$ on G includes an assignment of vertices to stations $s \in S$ including the configuration of fixed chargers and the scheduling of each mobile charger among all time slots.*

We use a utility function (see Eq. (16)) to evaluate the effectiveness of a charging plan. We set a limit on the budget to install fixed stations via Eq. 1.

$$\sum_{s \in S} x_i \cdot fee_i + \sum_{m \in M} i_{MC}(m) \cdot fee_{MC} \leq Budget. \quad (1)$$

Here, fee_i represents the installation cost of the i-th type of fixed charger, $i_{MC}(m)$ is a binary variable where $i_{MC}(m) =$

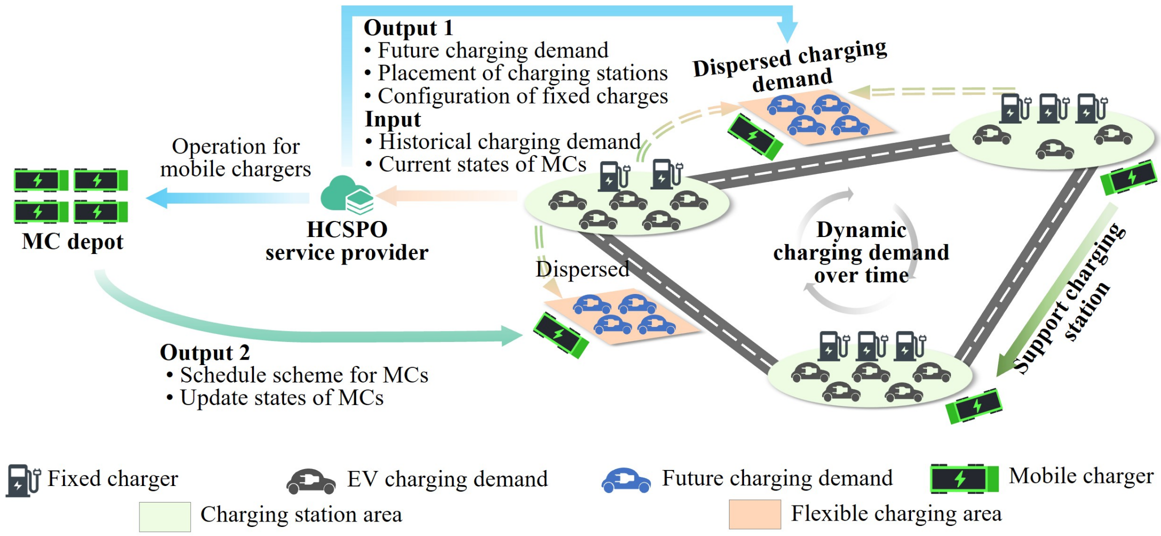


Figure 2: Overview of HCSPO

1 indicates that mobile charger m is employed; otherwise, it remains idle throughout the entire horizon. fee_{MC} denotes the operational cost of each mobile charger, and $Budget$ refers to the total budget allocated for the system.

We formalize HCSPO as follows. Given a road network $G = (V, E)$, time slots $T = 1, \ldots, |T|$, historical node-level demand, and perfect foresight of future demand, we seek a charging plan that satisfies all demand and maximizes a utility that balances social benefit against installation cost and queuing loss.

Figure 2 summarizes the workflow: before operation, the provider forecasts demand from historical data, selects station locations/configurations, and schedules mobile chargers to flexibly serve areas beyond fixed stations and ease peak-hour congestion.

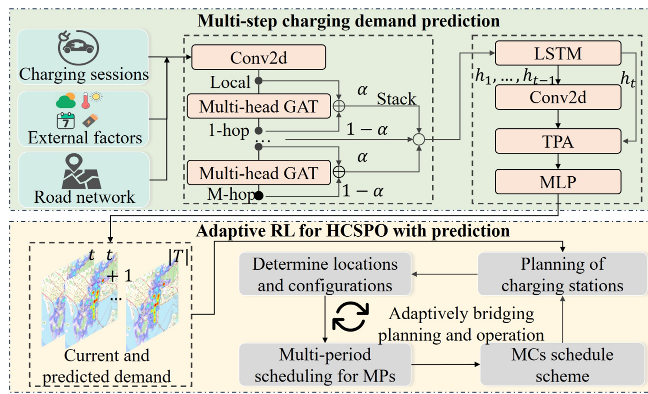


Figure 3: System framework

4 System Framework

This section outlines the HCSPO system framework (Figure 3). We begin with a multi-step charging demand prediction model based on historical data. Next, we adapt the utility model from [von Wahl *et al.*, 2022] into a rolling horizon framework that incorporates current and future demand for MPC-based decision-making. Finally, we propose two mobile charger scheduling strategies integrated with fixed station planning.

4.1 Multi-step Charging Demand Prediction

To anticipate future charging needs, we formulate a multi-step demand prediction problem as:

$$\hat{dem}^{t+1:t+\omega_{pred}} = f(dem^{t-\omega_{hist}+1:t}, F^{t-\omega_{hist}+1:t}), \quad (2)$$

where dem denotes historical node-level demand and F includes exogenous features (e.g., weather, day type). We adopt a spatiotemporal forecasting model inspired by [Qu *et al.*, 2024], which integrates a **Graph Embedding Module** and a **Multivariate Decoder**. The embedding module captures temporal trends via CNN and spatial correlations via GAT [Veličković *et al.*, 2018], enhanced by residual connections. The decoder uses TPA-LSTM [Shih *et al.*, 2019] to extract long-term temporal dependencies with multi-hop attention, enabling accurate multi-step forecasting over the graph. This predicted demand guides downstream optimization of CS and MC operations.

4.2 Utility Function

This section introduces a utility model to evaluate the effectiveness of a charging plan. Building on the utility model designed by [von Wahl *et al.*, 2022], we extend this model into a rolling horizon framework to assess the performance of a charging plan under time-varying demand. We also introduce a queuing loss metric that measures the number of EVs leaving the queue due to excessive wait times for recharging.

Benefit

The benefit function measures the coverage of the road network by charging infrastructures. Intuitively, nodes supported by a greater number of charging infrastructures suggest higher benefits. To quantify this, we define the influential radius of a charging station (CS) s as the distance within which it attracts EVs from nearby nodes, influenced by its charging supply capacity.

We first define the charging supply capacity as the total power provided by both fixed chargers and mobile chargers (if any) at time slot t, which can be calculated as follows:

$$C^t(s) = \sum_{i=1}^{n} x_i c_i + \sum_{m \in M} i_{MC}(m,s,t) c_{MC} k_{MC} \delta, \quad (3)$$

where c_i denotes the charging power provided by each fixed charger of type i, c_{MC} is the charging power provided by each mobile charger (MC), and $i_{MC}(m,s,t) : M \times V \times T \mapsto \{0,1\}$ is an indicator function that equals 1 if m is scheduled to s at time slot t, and 0 otherwise. $\delta \in [0,1]$ is the invalid time discount factor since MCs could not offer charging capacity while scheduling, which will be described in Section 4.3. Once the capacity of a CS s is determined, the influential radius of s at time slot t is: $R^t(s) = R_{\max} \frac{1}{1+\exp\left(-\tilde{C}^t(s)\right)}$, where $R_{\max}$ denotes the maximum influential radius (a model parameter), and $-\tilde{C}^t(s)$ represents the scaled capacity of s at t.

Next, we define the coverage of node v at t, $cov^t(v)$, as the set of stations whose influential radius can cover v: $cov^t(v) = \{s \in P \mid dist(v,s) < R^t(s)\}$, where $dist(v,s)$ is the distance between v and s. The benefit of the charging plan P is then formulated as follows:

$$benefit(P) = \frac{1}{|V||T|} \sum_{t \in T} \sum_{v \in V} \left(\sum_{i=1}^{|cov^t(v)|} \frac{1}{i} \right). \quad (4)$$

Cost

We evaluate the cost of a charging plan P in terms of $travel\ time$, $charging\ time$ and $waiting\ time$. However, since the presence of dynamic charging demand and MCs scheduling, EVs from the same node may visit different CSs at different time slots. Thus we also need to extend cost function into a rolling horizon context. In order to estimate $travel\ time$, we first define an indicator function $i_{CS}(v,s) : V \times S \mapsto \{0,1\}$ to be 1 if s is the CS assigned to v according to Station Seeking Algorithm in [Liu *et al.*, 2019], and 0 otherwise. Then the $travel\ time$ of P is:

$$travel(P) = \frac{1}{|T|} \sum_{t \in T} \sum_{v \in V} \sum_{s \in S} i_{CS}(v,s) \frac{dist(v,s)}{vel} dem^t(v), \quad (5)$$

where $vel = const$ is the average vehicle speed.

Next, we estimate the $charging\ time$ and $waiting\ time$ using the Pollaczek-Khintchine formula, as in [von Wahl *et al.*, 2022]. However, this approach assumes no limits on queue length, considering only the stability of the queuing system. In real-world applications, charging demand can vary over time, and demand may occasionally exceed station capacity. To address this, instead of enforcing stability constraints, we set a maximum average waiting time, W_{max}. Specifically, if the waiting time for a queue with length s exceeds $W_{\max}$, no new EV will join the queue. Let

$$\mu^t(s) = C^t(s)/B \quad (6)$$

be the service rate of s at t (where $B = const$ is the electricity required for recharging each EV). According to Pollaczek-Khintchine formula, estimated number of approaching EVs at t $D^t(s)$ is:

$$D^t(s) = \sum_{v \in V} \frac{i_{CS}(v,s)}{dist(v,s)} dem^t(v). \quad (7)$$

Then the average waiting time of station s with no stability constraints is:

$$W^t(s) = \frac{\rho^t(s)}{2\mu^t(s)\left(1-\rho^t(s)\right)}, \quad (8)$$

where

$$\rho^t(s) = D^t(s)/\mu^t(s) < 1 \quad (9)$$

is the stability constraints for queuing system in [von Wahl *et al.*, 2022; Liu *et al.*, 2019]. When we set maximum $W(s)$ to $W_{\max}$, and replace it with $W(s)$ into Eq. (6), (8) and (9), we can obtain the maximum $D^t_{\max}$ through:

$$D^t_{\max} = \frac{2W_{\max}\left(C^t(s)\right)^2}{\left(2W_{\max}C^t(s)+B\right)B}. \quad (10)$$

Therefore, we estimate the corrected number of EVs approaching CS s at t as follows:

$$\tilde{D}^t(s) = \begin{cases} D^t_{\max}, & D^t(s) \geq D^t_{\max} \\ D^t(s), & D^t(s) < D^t_{\max} \end{cases}, \quad (11)$$

and corrected $\tilde{W}^t(s)$ can be obtained by:

$$\tilde{W}^t(s) = \begin{cases} W_{\max}, & D^t(s) \geq D^t_{\max} \\ \frac{\rho^t(s)}{2\mu^t(s)(1-\rho^t(s))}, & D^t(s) < D^t_{\max} \end{cases}, \quad (12)$$

And we estimate $charging\ time$ of charging plan P through

$$charging(P) = \frac{1}{|T|} \sum_{t \in T} \sum_{s \in S} \frac{D^t(s)}{\mu^t(s)}. \quad (13)$$

Finally, we aggregate the travel time, charging time, and waiting time for a charging plan (P) into a single cost function:

$$\begin{aligned} cost(P) = {} & \alpha \cdot travel(P) \\ & + (1-\alpha) \cdot \left[waiting(P) + charging(P)\right], \end{aligned} \quad (14)$$

where $\alpha \in [0,1]$ is a weighting parameter.

Queuing Loss

Since we have introduced the maximum waiting time $W_{\max}$ and corrected estimated number of EVs $\tilde{D}^t(s)$ in Section 4.2, a metric so-called $queuing\ loss$ is introduced to measure the loss of those demand for avoiding long waiting time, i.e.,

$$queuing\ loss(P) = \sum_{t \in T} \sum_{s \in S} \max\{D^t(s) - \tilde{D}^t(s), 0\}. \quad (15)$$

Finally, we model utility function by combining three evaluation metrics through:

$$\begin{aligned} utility(P) = {} & \lambda_b \cdot benefit(P) + \lambda_c \cdot cost(P) \\ & + \lambda_q \cdot queuing\ loss(P), \end{aligned} \quad (16)$$

where λ_b, λ_c and λ_q are trade-off weights.

4.3 Mobile Chargers Scheduling

We propose two heuristic strategies for scheduling mobile chargers (MCs): supporting overloaded charging stations and establishing flexible charging areas. Additionally, we introduce policies for MC adjustment and recall to adapt to dynamic demand.

Supporting Charging Stations

To support overloaded charging stations, we employ a heuristic strategy: (1) identify the station s with the highest loss, calculated as $\lambda_c \cdot (waiting + charging) + \lambda_q \cdot queuing\ loss$; (2) sort all idle MCs by their distance to s; (3) select the nearest idle MC with sufficient energy ($> c_{MC}$) and not currently scheduled; (4) assign it to s and update the station's service capacity and related metrics accordingly.

Establishing Flexible Charging Areas

To supplement fixed charging stations, we employ a heuristic strategy: (1) identify the non-CS node v with the highest cumulative demand $dem^{t:|T|}(v)$; (2) sort all idle MCs by their distance to v; (3) select the nearest idle MC with sufficient energy ($> c_{MC}$) and not currently scheduled; (4) dispatch it to v to establish a temporary charging site without fixed infrastructure; (5) update system metrics such as capacity C, service rate μ, and adjusted demand $\tilde{D}$ accordingly.

Adjustment and Recall Policy for MCs

Relocating a MC m incurs a scheduling delay during which it cannot provide service. The next arrival time after scheduling is given by $\tau_m^{t+1} = \text{dist}(l_m^t, l_m^{t+1})/\text{vel}$. To account for the service loss during this period, we define a discount factor:

$$\delta_m^{t+1} = 1 - \min\left\{\frac{\tau_m^{t+1} - (t+1)H}{H}, 1\right\}, \tag{17}$$

where H is the duration of a time slot.

Since MCs have limited battery capacity, their remaining energy is updated as:

$$q_m^{t+1} = q_m^t - \sum_{s \in S} i_{MC}(m, s, t) \cdot \max\{c_{MC}\delta_m^t, 0\}. \tag{18}$$

If $q_m^t < c_{MC}$, the MC is recalled to a depot j for recharging. The required energy is $q = B_{MC} - q_m^t$, and the updated arrival time becomes $\tau_m^{t+1} = \text{dist}(l_m^t, j)/\text{vel} + C'/q$, where C' is the charging power.

5 Reinforcement Learning Framework

We propose an adaptive reinforcement learning framework (ARL-HCSPO) to jointly optimize station placement, fixed charger setup, and mobile charger scheduling, using demand prediction (Section 4.1) and utility modeling (Section 4.2) to shape the reward in dynamic environments.

Observation. The observation O^i at episode step i consists of two components: planning observation for CSs, denoted as O^i_{CS} and operation observation for MCs, denoted as $O^i_M C$. The planning observation is defined as follows $O^i_{CS} = \left\{\left(lon_v, lat_v, dem_v^{1,\cdots,|T|}, x\right)_v |\forall v \in V\right\}$, where lon_v and lat_v are the coordinate of node v, and $x = (x_1, \cdots, x_n)$ is the configuration of fixed chargers, which is only applicable when v is a CS.

Then we define the operation observation using equation $O^i_{MC} = \left\{\left(l_m^{1,\cdots,|T|}, q_m^{1,\cdots,|T|}, \tau_m^{1,\cdots,|T|}\right)_m |\forall m \in M\right\}$, which captures the location, arrival time after scheduling and remaining electricity of all time slots $t \in T$.

Action. We adopt a five-action space inspired by [von Wahl *et al.*, 2022], serving as neighborhood operations to modify the charging plan:

- *Create by Demand*: Add a CS at the node with highest demand using a precomputed capacity-cost lookup table over n charger types.
- *Create by Benefit*: Add a CS at the node with lowest coverage; configuration follows the same lookup rule.
- *Increase by Demand*: Add one charger to the node with highest demand.
- *Increase by Benefit*: Add one charger to a CS near the lowest-benefit node.
- *Relocate*: Move a charger from the lowest-benefit CS to the one with the highest unmet demand (e.g., waiting time or queuing loss).

To balance planning and real-time operation, after each action step i, we schedule mobile chargers (MCs) based on the updated plan. This coordination ensures flexibility for demand surges via MCs, while CSs provide stable long-term coverage.

Reward. Given the utility function described in Section 4.2, we define the reward function between transition as the difference between P^{i+1} and P^i, i.e., $r^i = utility(P^{i+1}) - utility(P^i)$.

6 Experiments

In this section, we will introduce how we perform our ARL-HCSPO method on a real-world application.

6.1 Datasets

We evaluate our approach on the **road network** of Nanshan District, Shenzhen, China, comprising 1663 nodes and 2964 edges, similar to [von Wahl *et al.*, 2022]. **Existing charging station** locations are collected from [Qu *et al.*, 2024] and used as a baseline. Due to the lack of precise charging demand data at the road level, we use charging session data from [Qu *et al.*, 2024] for **charging demand** estimation, allocating node-level demand via inverse distance weighting from each node to the nearest CS.

6.2 Evaluation Metric and Implementation

Based on the utility function in Section 4.2, we evaluate the effectiveness of the charging plan generated by our approach using several metrics: $benefit$, $cost$, $queuing\ loss$, $travel$, $charging$, and $waiting$.

In the utility model, we use the following parameters as default setting: $\alpha = 0.4$, $\lambda_b = 0.4$, $\lambda_c = 0.4$, $\lambda_q = 0.2$, $n = 3$, $k = 25$, $Budget = 5.4e7$ CNY, $E = 35$ kwh, $r_{\text{max}} = 2.5$ km, $H = 60$ min, T=3, $W_{\text{max}} = 30$ min, $|M| = 30$, $k_{MC} =$

10, $c_{\text{MC}} = 42$ kw, $B_{\text{MC}} = 105$ kwh. The rated charging power of fixed chargers, c_1, c_2 and c_3 are set to 7kw, 22kw and 50kw, respectively, with corresponding installation fees of 2700 CNY, 6750 CNY and 252000 CNY. The operation cost for each MC is 59400 CNY.

RL in this paper is applied on Stable Baselines 3 [Raffin *et al.*, 2021], implemented using DQN as basic model to train an optimal policy. We set learning rate to 0.01, buffer size to 10000, batch size to 128. The maximum number of training steps is set to 30000.

6.3 Baselines

To evaluate the effectiveness of our method (ARL-HCSPO), we compare it with several state-of-the-art baselines under the same budget constraint (Eq. 1). Since baselines do not utilize mobile chargers, they are allowed to deploy more fixed chargers for fairness.

- **EXISTING PLAN**: Uses the real-world deployment of location-fixed charging stations as a reference to quantify improvement.
- **HIGHEST DEMAND**: A greedy strategy that selects nodes with the highest cumulative demand until the budget is depleted; charger configurations follow our action space setup.
- **BOUNDING&OPTIMIZING+**: Based on [von Wahl *et al.*, 2022], this method enhances [Liu *et al.*, 2019] by integrating a configuration lookup table for greedy allocation.
- **PCRL**: A reinforcement learning-based approach from [von Wahl *et al.*, 2022] for optimizing fixed CS placement in real-world scenarios.

6.4 Experimental Results

In this experiment we apply our approach as well as other baselines to solve HCSPO problem on Nanshan district, Shenzhen datasets.

Evaluation

Table 1 presents the performance metrics of our approach compared to other baselines, where ARL-HCSPO consistently outperforms across all metrics. Notably, the $benefit$ metric sees a significant improvement at 205.1%, vastly exceeding other methods, indicating that our approach can substantially enhance the availability of charging infrastructure across the road network. Additionally, a 40.5% reduction in the $cost$ metric demonstrates our approach's ability to improve EV owners' charging satisfaction, especially reflected in $waiting$ metric with a 79.8% reduction.

For RL-based approach, i.e., PCRL and ARL-HCSPO, we initialize charging plan as empty before training. Figure 4 illustrates the training evolution of our approach compared to PCRL. As shown, our method (in blue) outperforms PCRL (in orange) by achieving higher episode rewards with more stable convergence.

Figure 5 depicts the distribution of charging stations generated by various methods. On this dataset, both the BOUNDING$OPTIMIZING+ and HIGH DEMAND approaches fail to show significant improvement over the EXISTING PLAN. In contrast, both PCRL and ARL-HCSPO exhibit noticeable enhancements, with ARL-HCSPO achieving the most extensive CS coverage, due to the integration of mobile charger scheduling. Moreover, ARL-HCSPO results in a more evenly distributed network of CSs across the road system.

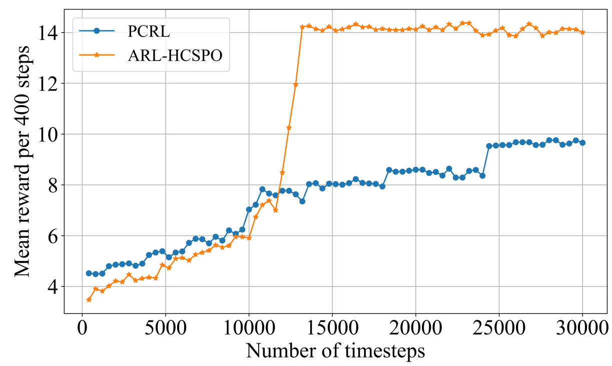


Figure 4: Training progression of ARL-HCSPO and PCRL, with mean rewards evaluated every 400 training steps.

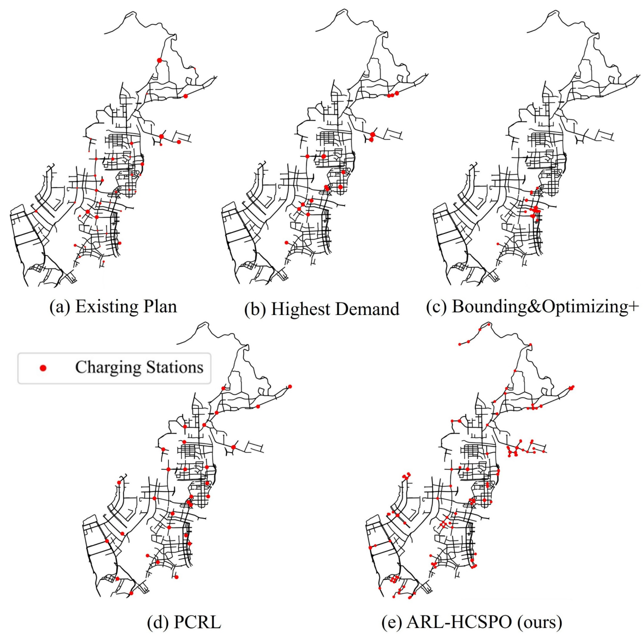


Figure 5: Charging plan comparison on Nanshan district, Shenzhen.

Ablation Study

To evaluate the effectiveness of each module within ARL-HCSPO, we compare our approach against three variants: (1) ARL-HCSPO (W/O MPC), which excludes the MPC policy; (2) ARL-HCSPO (W/O MCS1), which removes the supporting charging station scheduling strategy; and (3) ARL-HCSPO (W/O MCS2), which omits the flexible charging areas scheduling strategy.

The results in Table 2 indicate that all variants improve upon the EXISTING PLAN. Comparing our approach with ARL-HCSPO (W/O MPC), we observe a significant 82.3% increase in the $benefit$ metric, alongside modest gains in other metrics, underscoring the value of incorporating the

Approach	$benefit \uparrow$	$cost \downarrow$	$travel \downarrow$	$charging \downarrow$	$waiting \downarrow$	$queuing\ loss \downarrow$
EXISTING PLAN	100%	100%	100%	100%	100%	100%
HIGHEST DEMAND	93%	90%	91.9%	90.4%	72.7%	98.2%
BOUNDING&OPTIMIZING+	78%	97.3%	98.5%	123.2%	73.2%	95.7%
PCRL	126%	63.8%	71.1%	44.9%	21%	94.2%
ARL-HCSPO (ours)	**205.1%**	**59.5%**	**52.8%**	**43.3%**	**20.2%**	**86.7%**

Table 1: Results on Nanshan, Shenzhen dataset compared to baselines. Evaluation metrics where higher value are better are marked with ↑, otherwise are marked with ↓. Best results are marked bold.

Approach	$benefit \uparrow$	$cost \downarrow$	$travel \downarrow$	$charging \downarrow$	$waiting \downarrow$	$queuing\ loss \downarrow$
EXISTING PLAN	100%	100%	100%	100%	100%	100%
ARL-HCSPO (W/O MPC)	122.8%	62.7%	68.6%	45.4%	20.2%	91.9%
ARL-HCSPO (W/O MCS1)	**205.7%**	63.9%	70.2%	**42.6%**	26.2%	93%
ARL-HCSPO (W/O MCS2)	141.3%	**48.5%**	**29.1%**	45.3%	**11.9%**	**67.8%**
ARL-HCSPO	205.1%	59.5%	52.8%	43.3%	20.2%	86.7%

Table 2: Ablation study on Nanshan, Shenzhen dataset. Evaluation metrics where higher value are better are marked with ↑, otherwise are marked with ↓. Best results are marked bold.

demand prediction model and MPC policy. Additionally, the analysis of the two MC scheduling strategies reveals different focal points: ARL-HCSPO (W/O MCS1) emphasizes maximizing $benefit$, achieving a peak $benefit$ of 205.7%, slightly exceeding ARL-HCSPO (205.1%). Conversely, ARL-HCSPO (W/O MCS2) prioritizes cost-efficiency, recording the lowest $cost$ at 48.5%, with significant reductions in $travel$ (29.1%) and $waiting$ (11.9%).

The results from ARL-HCSPO demonstrate that a balanced approach, combining both strategies, can deliver a substantial $benefit$ increase of 63.8% with only an 11% increase in $cost$ compared to ARL-HCSPO (W/O MCS2).

Effect of K

The maximum number of fixed chargers, K, is a critical factor in determining the effectiveness of our approach. Figure 6 illustrates the impact of varying K values on different performance metrics when applying ARL-HCSPO. The results for all metrics are normalized by the scaled-down value of the EXISTING PLAN, as shown in Table 1.

For the $benefit$ metric, increasing K from 4 to 36 results in an optimal value of 244.4% at $K = 8$, after which there is a moderate decline. In contrast, metrics such as $cost$, $travel$, and $waiting$ continue to improve as K increases, reaching their optimal values of 51.6%, 38.3%, and 75.2%, respectively, at $K = 32$. The $queuing\ loss$ metric exhibits some fluctuation but generally trends downward.

In summary, while increasing K improves performance in terms of $cost$ and $queuing\ loss$, it is important to strike a balance by keeping K within a smaller range to optimize $benefit$. This balance is crucial for improving the overall supply of charging infrastructure.

Further sensitivity analysis could provide valuable managerial insights for future applications.

7 Conclusion

In conclusion, this paper presents a comprehensive solution to the HCSPO problem by integrating fixed charging station planning with dynamic mobile charger operation. We propose a demand prediction model and two heuristic scheduling strategies for mobile chargers, solved using an adaptive reinforcement learning algorithm. Extensive experiments on real-world datasets show that our method outperforms existing approaches across multiple metrics, offering an efficient solution to dynamic EV charging demand challenges.

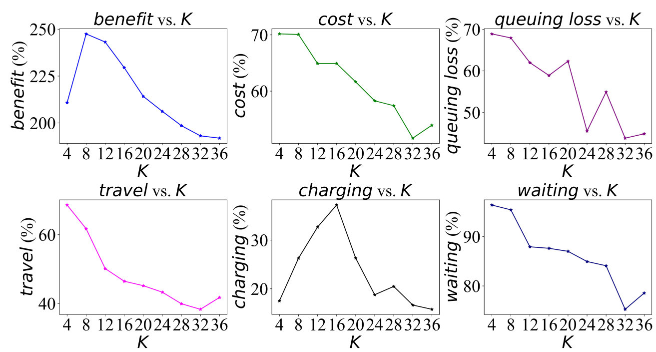


Figure 6: Impact of various metrics compared to EXISTING PLAN as percentages, evaluated against different values of K.

Acknowledgments

This work is mainly supported by the Guangdong Basic and Applied Basic Research Foundation (No. 2025A1515011994). This work is also supported by the National Natural Science Foundation of China (No. 62402414), Guangzhou Municipal Science and Technology Project (No. 2023A03J0011), the Guangzhou Industrial Information and Intelligent Key Laboratory Project (No. 2024A03J0628), and a grant from State Key Laboratory of Resources and Environmental Information System, and Guangdong Provincial Key Lab of Integrated Communication, Sensing and Computation for Ubiquitous Internet of Things (No. 2023B1212010007).

References

[Afshar *et al.*, 2021] Shahab Afshar, Pablo Macedo, Farog Mohamed, and Vahid Disfani. Mobile charging stations for electric vehicles — a review. *Renewable and Sustainable Energy Reviews*, 152:111654, 2021.

[Beyazıt and Taşcıkaraoğlu, 2023] Muhammed Ali Beyazıt and Akın Taşcıkaraoğlu. Electric vehicle charging through mobile charging station deployment in coupled distribution and transportation networks. *Sustainable Energy, Grids and Networks*, 35:101102, 2023.

[Cai *et al.*, 2014] Hua Cai, Xiaoping Jia, Anthony S. F. Chiu, Xiaojun Hu, and Ming Xu. Siting public electric vehicle charging stations in beijing using big-data informed travel patterns of the taxi fleet. *Transportation Research Part D: Transport and Environment*, 33:39–46, 2014.

[Hwang *et al.*, 2015] Seong Wook Hwang, Sang Jin Kweon, and Jose A. Ventura. Infrastructure development for alternative fuel vehicles on a highway road system. *Transportation Research Part E: Logistics and Transportation Review*, 77:170–183, 2015.

[Kchaou Boujelben and Gicquel, 2019] Mouna Kchaou Boujelben and Celine Gicquel. Efficient solution approaches for locating electric vehicle fast charging stations under driving range uncertainty. *Computers and Operations Research*, 109:288–299, 2019.

[Kchaou-Boujelben and Gicquel, 2020] Mouna Kchaou-Boujelben and Céline Gicquel. Locating electric vehicle charging stations under uncertain battery energy status and power consumption. *Computers and Industrial Engineering*, 149:106752–106752, 2020.

[Kchaou-Boujelben, 2021] Mouna Kchaou-Boujelben. Charging station location problem: A comprehensive review on models and solution approaches. *Transportation Research Part C: Emerging Technologies*, 132, 2021.

[Kong *et al.*, 2019] Weiwei Kong, Yugong Luo, Guixuan Feng, Keqiang Li, and Huei Peng. Optimal location planning method of fast charging station for electric vehicles considering operators, drivers, vehicles, traffic flow and power grid. *Energy*, 186, 2019.

[Lee and Han, 2017] Chungmok Lee and Jinil Han. Benders-and-price approach for electric vehicle charging station location problem under probabilistic travel range. *Transportation Research Part B: Methodological*, 106:130–152, 2017.

[Li *et al.*, 2017] Meng Li, Yinghao Jia, Zuojun Shen, and Fang He. Improving the electrification rate of the vehicle miles traveled in beijing: A data-driven approach. *Transportation Research Part A: Policy and Practice*, 97:106–120, 2017.

[Li *et al.*, 2024] Hua Li, Dongmin Son, and Bongju Jeong. Electric vehicle charging scheduling with mobile charging stations. *Journal of Cleaner Production*, 434:140162, 2024.

[Liu *et al.*, 2019] Qiyu Liu, Yuxiang Zeng, Lei Chen, and Xiuwen Zheng. Social-aware optimal electric vehicle charger deployment on road network. In *Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, SIGSPATIAL '19, page 398–407, New York, NY, USA, 2019. Association for Computing Machinery.

[Liu *et al.*, 2024] Linfeng Liu, Su Liu, Jiagao Wu, and Jia Xu. A placement strategy for idle mobile charging stations in ioev: From the view of charging demand force. *IEEE Transactions on Intelligent Transportation Systems*, 25(5):3870–3884, 2024.

[Qu *et al.*, 2024] Haohao Qu, Haoxuan Kuang, Qiuxuan Wang, Jun Li, and Linlin You. A physics-informed and attention-based graph learning approach for regional electric vehicle charging demand prediction. *IEEE Transactions on Intelligent Transportation Systems*, pages 1–14, 2024.

[Raffin *et al.*, 2021] Antonin Raffin, Ashley Hill, Adam Gleave, Anssi Kanervisto, Maximilian Ernestus, and Noah Dormann. Stable-baselines3: Reliable reinforcement learning implementations. *Journal of Machine Learning Research*, 22(268):1–8, 2021.

[Shahraki *et al.*, 2015] Narges Shahraki, Hua Cai, Metin Turkay, and Ming Xu. Optimal locations of electric public charging stations using real world vehicle travel patterns. *Transportation Research Part D: Transport and Environment*, 41:165–176, 2015.

[Shih *et al.*, 2019] Shun-Yao Shih, Fan-Keng Sun, and Hung yi Lee. Temporal pattern attention for multivariate time series forecasting, 2019.

[Tang *et al.*, 2020] Peng Tang, Fang He, Xi Lin, and Meng Li. Online-to-offline mobile charging system for electric vehicles: Strategic planning and online operation. *Transportation Research Part D: Transport and Environment*, 87:102522, 2020.

[Ting *et al.*, 2024] Lo Pang-Yun Ting, Chi-Chun Lin, Shih-Hsun Lin, Yu-Lin Chu, and Kun-Ta Chuang. Multi-agent reinforcement learning for online placement of mobile ev charging stations. In De-Nian Yang, Xing Xie, Vincent S. Tseng, Jian Pei, Jen-Wei Huang, and Jerry Chun-Wei Lin, editors, *Advances in Knowledge Discovery and Data Mining*, pages 284–296, Singapore, 2024. Springer Nature Singapore.

[Tu *et al.*, 2016] Wei Tu, Qingquan Li, Zhixiang Fang, Shih lung Shaw, Baoding Zhou, and Xiaomeng Chang. Optimizing the locations of electric taxi charging stations: A spatial–temporal demand coverage approach. *Transportation Research Part C: Emerging Technologies*, 65:172–189, 2016.

[Veličković *et al.*, 2018] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks, 2018.

[von Wahl *et al.*, 2022] Leonie von Wahl, Nicolas Tempelmeier, Ashutosh Sao, and Elena Demidova. Reinforcement learning-based placement of charging stations in urban road networks. In *Proceedings of the 28th ACM*

SIGKDD Conference on Knowledge Discovery and Data Mining, KDD '22, page 3992–4000, New York, NY, USA, 2022. Association for Computing Machinery.

[von Wahl *et al.*, 2023] Leonie von Wahl, Clara Mautsch, Stefan Ladwig, Michael Nolting, and Nicolas Tempelmeier. Towards data-driven planning of mobile charging infrastructure for electric vehicles. In *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Sustainable Mobility*, SuMob '23, page 13–22, New York, NY, USA, 2023. Association for Computing Machinery.

[Wang *et al.*, 2023] Shiqi Wang, Yuze Li, Anthony Chen, and Chengxiang Zhuge. Electrification of a citywide bus network: A data-driven micro-simulation approach. *Transportation Research Part D: Transport and Environment*, 116:103644–103644, 2023.

[Xiang *et al.*, 2019] Yue Xiang, Zhuozhen Jiang, Chenghong Gu, Fei Teng, Xiangyu Wei, and Yang Wang. Electric vehicle charging in smart grid: A spatial-temporal simulation method. *Energy*, 189, 2019.

[Yang *et al.*, 2017] Jie Yang, Jing Dong, and Liang Hu. A data-driven optimization-based approach for siting and sizing of electric taxi charging stations. *Transportation Research Part C: Emerging Technologies*, 77:462–477, 2017.

[Yin *et al.*, 2023] Wanjun Yin, Jianbo Ji, and Xuan Qin. Study on optimal configuration of ev charging stations based on second-order cone. *Energy*, 284, 2023.