TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214
DOI: 10.26599/TST.2025.9010059

dRefine: A Data-Driven Deep Reinforcement Learning Model for Battery Swapping Station Network Refinement

Han Lu, Wenhui Cheng, Ruixue Huang, Chaocan Xiang*, Xiao Zheng, Qing Zhou, and Jianchao Zheng

Abstract: In recent years, battery swapping services have experienced rapid growth, positioning themselves as a significant alternative to traditional charging stations. One determinant of their success lies in the ability to continuously refine the battery swapping station network, *i.e.,* redistributing batteries across stations, to meet the fluctuating demands of electric scooters. However, achieving such network refining is far from straightforward and presents two major challenges. First, for regions where stations have been recently deployed, the lack of historical data makes it difficult to accurately predict user demand. Secondly, effective battery scheduling is inherently complex, as it requires meticulous coordination across all stations to dynamically balance users' fluctuating long-term demand. To tackle these challenges, we propose *dRefine*, a novel data-driven battery swapping station network refinement system. Specifically, to address the first challenge, we develop a station-level spatiotemporal representation-guided conditional diffusion model, which leverages data from regions with established networks to predict demand in regions lacking historical data. For the second challenge, we develop a demand-oriented deep reinforcement learning model that dynamically optimizes battery scheduling strategies. By continuously learning from real-time demand patterns and operational feedback, it ensures efficient and adaptive battery redistribution across the entire network. We evaluate *dRefine* using a real-world dataset encompassing 388 stations, 41,358 batteries, and 108,574 electric scooter users. Extensive experimental results demonstrate that our method consistently outperforms state-of-the-art approaches by an average of 29.28%.

Key words: battery swapping station; deep reinforcement learning; diffusion model

1 Introduction

In recent years, battery swapping services have witnessed a worldwide prevalence[2]. It is reported that the global battery swapping market was valued at approximately $1.7 billion in 2022 and is projected to grow to $11.8 billion by 2027[3]. Under the battery swapping mode, electric scooters (e-scooters) can swap their depleted batteries for fully charged ones within minutes — significantly faster than conventional recharging via charging posts. This offers great convenience for e-scooters and effectively alleviates their range anxiety [4–7], while also contributing to

• A preliminary conference version of this paper was honored with the Best Paper Award at the 17th China Conference on Internet of Things (Wireless Sensor Network) (CWSN 2023)[1].

• Han Lu, Wenhui Cheng, Ruixue Huang, Chaocan Xiang, and Qing Zhou are with the College of Computer Science, Chongqing University, Chongqing 400044, China. E-mail: {luhan, huangruixue}@stu.cqu.edu.cn; {xiangchaocan, wenhuicheng, tzhou}@cqu.edu.cn.

• Xiao Zheng is with the School of Computer Science and Technology, Anhui University of Technology, Ma'anshan 243032, China. E-mail: xzheng@ahut.edu.cn.

• Jianchao Zheng is with the Academy of Military Science, Beijing 100091, China. E-mail: longxingren.zjc.s@163.com.

∗ To whom correspondence should be addressed (Chaocan Xiang, xiangchaocan@cqu.edu.cn).

Manuscript received: 2025-01-24; revised: 2025-02-26; accepted: 2025-04-02

reduced fossil fuel consumption and lower greenhouse gas emissions[8].

However, despite the significant benefits of the Battery Swapping Station (BSS) system, how to operate these BSS networks effectively remains a major challenge. The inherent dynamics of user swapping behavior, coupled with the complexities of urban environments, frequently lead to significant imbalances between battery supply and user demand (a detailed examination of this phenomenon is presented in Fig. 3). Therefore, refining the BSS network to dynamically meet users' battery swapping demand is a critical issue that urgently needs to be addressed. To this end, this paper delves into the BSS network refinement problem, focusing on the dynamic reallocation of batteries across urban regions to meet the long-term and fluctuating user demand for battery swapping. This problem involves the following challenges:

- *Accurately predicting user demand for all urban regions is challenging due to the lack of historical data in some regions with newly deployed stations.* Predicting users' battery swapping demand at each station is a fundamental step in optimizing battery allocation across the network. Historical operational data from stations can help to achieve accurate predictions of demand. However, such data is not always accessible for all regions. For example, in areas where battery swapping stations have been newly deployed, insufficient historical data makes it challenging to forecast the user demand precisely.

- *Effectively scheduling batteries across regions is complex because it requires coordinated decision-making among all stations to meet users' dynamic and long-term demands.* Battery allocation must consider the long-term user demand for battery swapping, which is continuously evolving due to dynamic environmental factors. Furthermore, each station requires a unique scheduling strategy, and these strategies are interdependent, influencing one another across the network. As a result, designing a dynamic and adaptive battery scheduling method poses a significant and complex challenge.

To address the challenges above, we propose *dRefine*, a data-driven battery swapping station network Refinement system. *dRefine* is structured as a two-stage process comprising: a prediction stage and an optimization stage. Specifically, in the prediction stage, we propose a novel conditional diffusion model enhanced with station-level spatiotemporal representations. This model effectively transfers knowledge from data-rich regions with established stations to accurately forecast user demand in data-scarce regions. In the optimization stage, we propose a demand-driven deep reinforcement learning model to globally optimize battery relocation decisions for each battery swapping station, aiming to maximize station operational benefits by effectively meeting long-term and dynamic user demand for battery swapping.

To summarize, we make the following three contributions:

- To the best of our knowledge, we are the first to explore the BSS network refinement problem using real-world operational data from battery swapping systems. Our research can effectively alleviate the prevalent user demand-supply imbalance problem and provide valuable insights for the optimization of BSS networks.

- We design a data-driven BSS network refinement system consisting of two main components: (i) a station-level spatiotemporal representation-guided conditional diffusion model to predict user demand in data-scarce regions, and (ii) a demand-oriented deep reinforcement learning model that optimizes battery allocation across regions to meet users' evolving long-term demands.

- We evaluate *dRefine* through extensive experiments on real-world battery swapping datasets from four major Chinese cities: Beijing, Shanghai, Chengdu, and Nanjing, covering 388 BSSs, 41,358 batteries, and 108,574 users. The results show that *dRefine* improves the operational benefit of BSS networks by an average of 29.28%, compared to existing state-of-the-art methods.

The rest of this paper is organized as follows: Sec. 2 provides a review of related work. Sec. 3 presents the motivation for this research, supported by large-scale dataset analyses. Sec. 4 formulates the research problem. In Sec. 5, we detail the novel data-driven battery swapping station network refinement system. The evaluation results are presented in Sec. 6. Finally, Sec. 7 concludes the paper.

2 Related Work

2.1 Charging and swapping station optimization

Electric scooters and electric vehicles have rapidly taken over the transportation market due to their green advantages[9–12]. With the expansion of the scale of electric transportation, studies on the optimization of charging and swapping stations have increased[13–15]. Wang *et al.* [16] propose a usage balancing method to schedule bikes among stations by considering the quality of service and bike maintenance. Liu *et al.* [17] predict station demand and plan routes, addressing the rebalancing problem in bike sharing systems. Wu *et al.* [18] develop a battery swapping station model to minimize costs by optimizing the charging schedule for swapped batteries. Du *et al.* [19] optimize the location and size of electric vehicle chargers, proposing an approximation algorithm with a theoretical bound. Lv *et al.* [20] develop a genetic algorithm based on a BSS localization model to minimize cost. Su *et al.* [21] use an adaptive genetic algorithm to minimize the peak-to-valley load difference. Tao *et al.* [22] consider battery life and optimize charging infrastructure to reduce overdischarge rate. Zheng *et al.* [23] present an optimal BSSs distribution framework based on life cycle cost to control reinforcement costs. Anirudh *et al.* [24] propose an e-bike station management framework to maximize the charge level of e-bikes and meet user demand by employing the deep deterministic policy gradient algorithm. Unlike all the above methods, our model considers the distribution of geographic information, uses demands and BSSs to adapt to the changing swapping environment, and dynamically redistributes batteries across BSSs.

2.2 Charging and swapping demand prediction

Charging and swapping demand are important factors in the optimization of energy facilities[25–27]. Zhao *et al.* [28] propose a simulation model to estimate battery swapping demand using data from shared e-scooter services. Recently, He *et al.* [29] provide the time-space probability distribution of swapping electric vehicles based on trip characteristics. Zhang *et al.* [30] simulate the behavior of electric vehicle users and uncertain highway traffic conditions based on Stochastic Dynamic UE. Sarker *et al.* [31] define battery demand in a range and apply robust inventory theory to consider the uncertainty of battery demand. Wang *et al.* [32] develop probability functions for the parking duration and State of Charge of electric freight vehicles to model charging demand. However, these studies rely mainly on historical demand data from urban regions, overlooking the issue of some regions lacking historical data. This paper proposes a demand prediction method based on a diffusion model, using data from regions with established networks to predict demand in data-scarce regions.

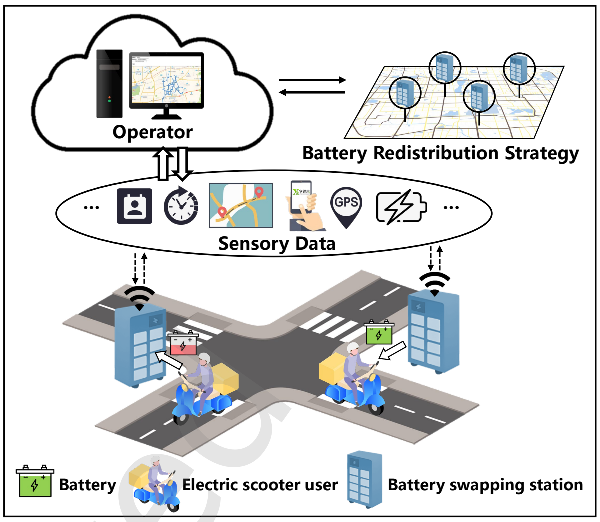


Fig. 1 The framework for the BSS network, consisting of BSSs, shared batteries, users, and operators.

3 Motivation

In this section, we first present the workflow of the battery swapping network. Next, we introduce a real-world dataset of battery swapping operations. Based on this dataset, we conduct a comprehensive data-driven analysis and derive novel insights.

3.1 BSS network operational workflow

As presented in Fig. 1, a BSS network typically consists of four main components: BSSs, shared batteries, users, and operators. BSSs store and charge batteries while utilizing sensors to monitor their status. Batteries are equipped with multiple sensors that periodically capture location and status information, which is then uploaded to the operator. Users exchange their discharged batteries for fully charged ones at any station of the network, where a swapping record is generated. The discharged batteries are then recharged at the BSSs until picked up by other users. The operator collects and monitors large-scale sensor data from swapping stations, users, and shared batteries, managing battery allocation in response to user swapping requests.

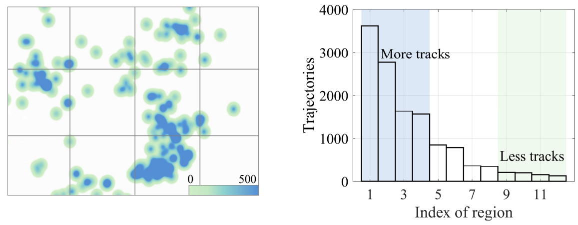
In particular, if each station in the BSS network has an unlimited supply of shared batteries, users could easily obtain a fully charged battery at any time. However, due to the high cost, operators typically rely on experience to allocate a limited number of batteries to each BSS. This imbalance results in some stations frequently facing battery shortages while others have a surplus, ultimately impacting the system's profitability. Therefore, rebalancing the distribution of batteries among swapping stations is essential. To evaluate the opportunities and feasibility of battery rebalancing, we perform a series of data analyses on a large-scale real-world dataset.

3.2 Large-scale dataset collection and analysis

Data collection. We collaborate with Xianglilai Technology Co., Ltd.[33], which operates BSS networks across 42 cities with 388 battery swapping stations. We obtain a large-scale real-world dataset from these networks, spanning six months (*i.e.,* from January 2024 to June 2024). The dataset includes sensor data and transaction records for 41,358 shared batteries and 108,574 users. To protect user privacy, we replace personal identifiers (*e.g.,* user IDs) with anonymous identifiers and add noise to geographic location data. The dataset comprises three categories of data: user swapping data, battery data, and battery swapping station data. The detailed descriptions of these data types are as follows:

- *User swapping data.* It includes the battery sharing usage records for all users. Each usage record contains details such as the order number, user ID, the time and station of the battery swap, the battery ID returned by the user, and the battery ID picked up by the user, etc.
- *Battery data.* The data is collected and uploaded every 15 minutes via sensor devices integrated within the battery. Each record includes a timestamp, battery ID, State of Charge (SoC), and the battery's latitude and longitude coordinates.
- *BSS data.* It describes the status information for each BSS, including key details such as swapping station ID, swapping station name, the latitude and longitude of the swapping station, and the number of batteries in the station.

Data analysis. Based on the real-world dataset, we conduct extensive data analysis to highlight the importance of optimizing BSS networks.



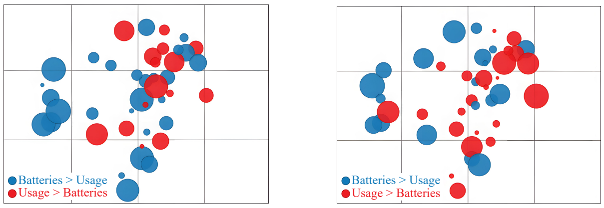
(a) Spatial distribution of e-scooter trajectories (b) E-scooter trajectory count in regions

Fig. 2 (a) Spatial distribution of e-scooter trajectories and (b) e-scooter trajectory count in regions.

(1) *Concentrated e-scooter trajectories distribution.* We first visualize the trajectories of e-scooter users in urban regions using battery data from a month, as shown in Fig.2a. In this visualization, areas marked in blue indicate higher user activity, reflecting a greater density of trajectories, while green areas represent lower activity levels. In Fig.2b, we present the regional statistics of the trajectory data, indicating that user trajectories are highly concentrated in some regions, with over 83% of the trajectories confined to just 28% of the road networks across the four cities. This pattern arises because users typically travel between fixed origins and destinations based on their preferences and mobility habits. Such spatial concentration can lead to overcrowding at certain BSSs, while others remain underutilized for extended periods.

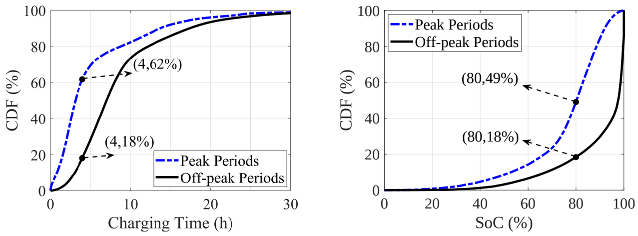
(2) *Unbalanced battery supply and usage.* Based on the findings from the user trajectory distribution, we analyze the user flow and battery distribution during two time periods (10 am ˜11 am and 5 pm ˜ 6 pm) within the BSS network, as shown in Fig. 3. Each circle represents a BSS, with its size reflecting the absolute difference between user flow and the number of available batteries. Red circles denote stations with few available batteries but high user flow, while blue dots represent stations with abundant batteries but low user flow. We can find that the supply and demand for batteries are geographically and temporally imbalanced. For example, some BSSs experience higher demand than supply from 10 am to 11 am (with more users than available batteries), while the supply exceeds demand from 5 pm to 6 pm.

(3) *Impact on the SoCs of swapped batteries.* Furthermore, we analyze the impact of imbalance on SoCs of the swapped-out batteries. Fig. 4a shows the Cumulative Distribution Function (CDF) of single charging times of batteries at a station during peak periods (10:00 am ˜ 11:00 am) and off-peak periods



(a) 10:00 am ˜ 11:00 am (b) 5:00 pm ˜ 6:00 pm

Fig. 3 Unbalanced battery supply and usage of BSS network during different time periods.



(a) Charging times for batteries (b) SoCs of picked-up batteries

Fig. 4 (a) Charging times for batteries and (b) SoCs of picked-up batteries.

(3:00 pm ˜ 4:00 pm). At peak periods, 62% of batteries are charged for less than 4 hours, whereas only 18% of batteries at off-peak periods are charged for less than 4 hours. This discrepancy arises because the user flow at the station during peak hours significantly exceeds the availability of fully charged batteries. The frequent pick-up and return of batteries during these times result in insufficient charging durations for most batteries. Fig. 4b presents the CDF of battery SoCs when they are picked up during peak and off-peak periods. Approximately 49% of batteries picked up during peak periods have an SoC below 80%, compared to only 18% during off-peak periods. This significant decrease in SoCs during peak periods contributes to range anxiety among e-scooter users.

In summary, the analysis results indicate that *the battery swapping station network suffers from a severe imbalance in battery supply and usage. Effectively reallocating battery distribution across stations to meet dynamic user demand is an urgent and crucial task.*

4 Problem formulation

BSS network model. Before presenting our research problem, we first provide a detailed introduction to the model of the BSS network:

- *BSS network:* The BSS network is represented as a graph $G = (V, E)$, where $V = \{v_1, v_2, \dots, v_I\}$ is the set of nodes, with each node v_i corresponding to a BSS. The edge set E contains edges $e_{ij} = (v_i, v_j) \in E$, each of which is associated with a feature vector that encodes the geographical distance between BSSs.

- *User demand:* We denote the number of users visiting BSS v_i during time slot t as the user demand, represented by $d_{t,i}$. The user demand sequences for the set of BSSs are then given by $D = \{D_1, D_2, \dots, D_T\}$, where $D_t = \{d_{t,1}, d_{t,2}, \dots, d_{t,I}\}$. In this paper, the time slot granularity is defined in hours.

- *Multi-source environment features:* The multi-source environment features for BSS v_i during time slot t are represented by $o_{t,i}$, which consists of road networks, population density, and Point of Interests (POIs). The multi-source environment features for the set of BSSs are given by $O = \{O_1, O_2, \dots, O_T\}$, where $O_t = \{o_{t,1}, o_{t,2}, \dots, o_{t,I}\}$.

- *Battery distribution:* We denote the capacity of each BSS v_i (*i.e.*, the number of batteries) during time slot t as $n_{t,i}$. The battery distribution for the set of BSSs is given by $N = \{N_1, N_2, \dots, N_T\}$, where $N_t = \{n_{t,1}, n_{t,2}, \dots, n_{t,I}\}$.

Based on the above BSS network model, we formulate the research problem as follows:

BSS Network Refinement Problem (BNR). Given the maximum capacity $n^m_{t,i}$ of the BSS v_i, the maximum number of batteries n^M that can be scheduled in the BSS network, and the number of batteries relocated from BSS i to BSS j, denoted as $n_{t,i,j}$, we optimize the number of batteries $n_{t,i}$ in each BSS to maximize the benefits of the BSS network, which is mathematically formulated as follows:

$$\arg\max_{n_{t,i}} \sum_{t=T+1}^{T+T'} \sum_{i=1}^{I} \min\left(d_{t,i}, n_{t,i}\right) + o_{t,i}, \tag{1}$$

$$\text{s.t.} \quad 1 \le n_{t,i} \le n^m_{t,i} \ \forall t, \forall i, \tag{2}$$

$$n_{t,i} = n_{t-1,i} + \sum_{j=1}^{I} n_{t,j,i} - \sum_{j=1}^{I} n_{t,i,j}, \tag{3}$$

$$\sum_{i=1}^{I} \sum_{j=1}^{I} (n_{t,i,j} - n_{t,j,i}) \le n^M \ \forall t, \forall i. \tag{4}$$

The objective function (1) aims to maximize the benefits of satisfying predicted user demand and the potential social benefits. These social benefits are quantified through multi-source environmental features. For instance, when an urban region develops more POIs, it has the potential to attract a greater number of new users. Constraint (2) limits the number of batteries at BSS v_i from 1 to maximum capacity $n_{t,i}^{m}$ to ensure connectivity between BSSs. Constraint (3) indicates that the battery count for BSS v_i at t is the battery count at $t-1$ plus the number of batteries scheduled to BSS v_i at t. Constraint (4) limits the total number of relocated batteries in the BSS network to a maximum of n^M to control scheduling costs.

5 Methods

To solve the BNR problem, we propose *dRefine*, a data-driven battery swapping station network refinement system, which is shown in Fig. 5. It is composed of two key components: diffusion model-based cross-region demand prediction and deep reinforcement learning-based battery redistribution.

- *Diffusion model-based cross-region demand prediction*: For regions where stations have been recently deployed, predicting future swapping demand based on historical data becomes infeasible. Fortunately, recent advances in conditional diffusion models present a new opportunity. These models can generate the swapping demand for a specific time and BSS based on spatiotemporal contextual information, without relying on historical data. Thus, we formulate the learning of the spatiotemporal pattern of battery swapping demand as a denoising diffusion process. Specifically, BSS network and multi-source environmental data are integrated as spatial information, while time slot data represent temporal information. They are fed into a spatiotemporal condition encoder, which extracts and fuses spatial and temporal features to produce the spatiotemporal context representations for BSSs. These representations then guide the conditional diffusion model to predict swapping demand sequences for BSSs.

- *Deep reinforcement learning-based battery redistribution*: Traditional supervised learning methods struggle to capture and model dynamics of user demand, supply and demand situation, and swapping environment, making it challenging to design an effective dynamic refinement strategy for the BSS network. Deep Reinforcement Learning (DRL) can learn the spatial effects and long-term temporal effects of the relocation strategy for the supply-demand balance. So we form the refinement strategy for BSSs as a complex sequential decision-making problem, aligning with the characteristics of the Multi-agent Markov Decision Processes (MMDPs)[34], and learn it through a multi-agent DRL. Specifically, the BSS agents, based on the observed state at each moment (including swapping demand, battery distribution, and environmental information), transfer batteries between BSSs to maximize the benefits of satisfying predicted demand and the potential social benefits.

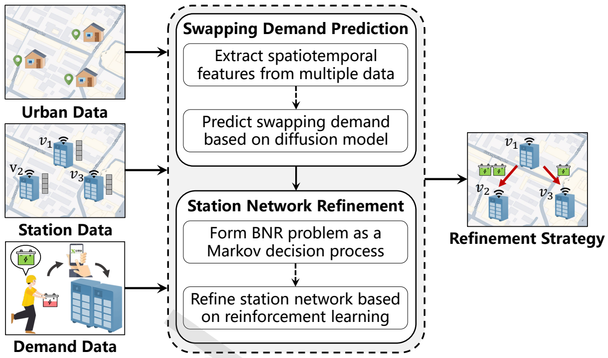


Fig. 5 Overview of *dRefine*, a novel data-driven battery swapping station network refinement system.

5.1 Diffusion model-based cross-region demand prediction

Inspired by previous work[35], we propose a station-level spatiotemporal representation-guided conditional diffusion model in this section, which learns the relationship between the spatiotemporal features of BSSs and user demand patterns. As shown in Fig. 6, it takes spatial and temporal information into a spatio-temporal condition encoder (see Sec. 5.1.1) to form conditions that guide the forward diffusion and reverse denoising processes (see Sec. 5.1.2) for predicting swapping demand.

5.1.1 Station-level spatiotemporal representation

To obtain generalized spatiotemporal context representations for BSSs, we design a spatiotemporal condition encoder. Specifically, we use a Multilayer Perceptron (MLP) to extract temporal features and a Graph Convolutional Network (GCN) to capture spatial features. A self-attention mechanism fuses features to generate the generalized context representation.

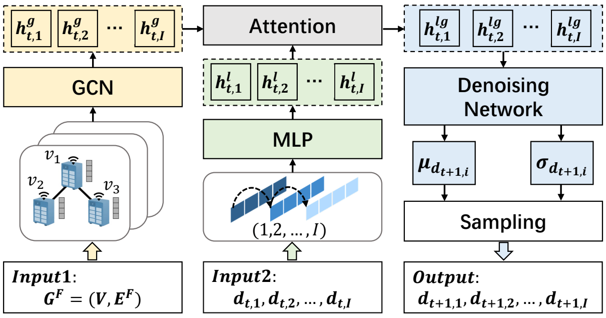


Fig. 6 The framework of station-level spatiotemporal representation-guided conditional diffusion model.

Temporal representation. To explore the temporal features, we conduct a series of data analyses. As shown in Fig. 7, we visualize user swapping demand in an urban area of Chengdu over a two-week period. We can observe a significant daily and weekly pattern, with three peak demand periods each day (10 ˜ 11 am, 5 ˜ 6 pm, and 8 ˜ 9 pm), and the weekday demand is generally higher than the weekend demand. In addition, we segment the weather conditions into different levels: W1 (Sunny, Cloudy), W2 (Foggy), W3 (Rainy, Snowy), and W4 (Heavy Rainy, Heavy Snowy) in Fig. 8a. We then analyze the impact of weather conditions on user swapping demand. Statistically, it is a common trend for the swapping demand to decrease during adverse weather conditions. Therefore, we extract the day of the week, hour of the day, and weather of the day as temporal features. These features are discretized into one-hot vectors and concatenated as input to an MLP, obtaining the final temporal representation of BSS v_i at t, denoted as $h^l_{t,i}$:

$$h^l_{t,i} = MLP\left([h^d_{t,i}, h^o_{t,i}, h^w_{t,i}]\right), \tag{5}$$

where $h^d_{t,i}, h^o_{t,i}, h^w_{t,i}$ represent the encoded information in terms of the day of the week, hour of the day, and weather of the day.

Spatial representation. We observe that under the same time and weather conditions, there are large differences in swapping demand patterns across BSSs in different regions. For example, Fig. 8b shows the swapping demand of BSSs in residential and commercial regions. We find that in residential region, swapping demand peaks at 8:00 am ˜ 9:00 am, while in commercial region, it peaks at 6:00 pm ˜ 7:00 pm. Therefore, the swapping demand of a BSS is related to the type of the region it located at. We can use the functional similarities between regions to enable mutual learning of cross-regional spatial representations. Since the functionalities of regions are related to multi-source environmental factors[36–38], we measure them based on the types and quantities of environmental factors. Specifically, we define $o^u_{t,i}$ as the number of environmental factors of type u covered by BSS v_i at time t. Therefore, the functional vector of BSS v_i is given by $o_{t,i} = \{o^1_{t,i}, o^2_{t,i}, \ldots, o^U_{t,i}\}$, where U is the total number of environmental factor types. We construct a functional similarity graph $G^F = (V, E^F)$, where each node represents a BSS v_i, and each edge $e^F_{ij} \in E^F$ is defined by the cosine similarity between $o_{t,i}$ and $o_{t,j}$, as follows:

$$e^F_{ij} = \frac{o_{t,i} o_{t,j}}{\|o_{t,i}\| \|o_{t,j}\|}. \tag{6}$$

We adopt GCN to capture the spatial dependency of BSSs. The output at the z-th layer of the GCN can be obtained as

$$H^{g,(z)}_t = f\left(\hat{E}^F H^{g,(z-1)}_t W^{(z-1)}\right), \tag{7}$$

where f denotes the activation function, $\hat{E}^F$ represents the normalized of E^F, and $W^{(z-1)}$ is the weight matrix. Thus, the spatial embedding of the BSS v_i at t is denoted as $h^g_{t,i}$, from the output of the GCN.

Attention-based representation fusion. Specifically, we simultaneously learn spatial attention and temporal attention to capture their fine-grained interactions adaptively. We project the vectors to get the query vector $Q_i = H^i W^Q$, the key vector $K_i = H^i W^K$, and the value vector $V_i = H^i W^V$. H^i is the embedding ($i \in l, g$). W^Q, W^K, and W^V are the weights of linear projections. We then calculate the importance of spatial and temporal representations as

$$u_i = softmax\left(\frac{QK_i^T}{\sqrt{d}}\right), i \in l, g, \tag{8}$$

where d is the dimension of the Q and K_i vectors. The spatiotemporal representation H^{lg} is calculated as

$$H^{lg} = (u_l V_l + u_g V_g) W^O, \tag{9}$$

where W^O is the output projection matrix.

Algorithm 1: Training for each swapping demand $d_{t,i}$

Input: $d_{t,i}$; $h^{lg}_{t,i}$;
Repeat
  $d^0_{t,i} \sim q(d^0_{t,i})$;
  $k \sim \text{Uniform}(1, 2, \ldots, K)$;
  $\epsilon \sim \mathcal{N}(0, I)$;
  Take gradient descent step on
  $\nabla_\theta \|\epsilon - \epsilon_\theta(\sqrt{\bar{\alpha}_k} d^0_{t,i} + \sqrt{1-\bar{\alpha}_k}\epsilon, h^{lg}_{t,i})\|^2$;
Until converged

5.1.2 Spatiotemporal-guided diffusion model

We transform demand data into pure noise by gradually adding noise and then learn the reverse

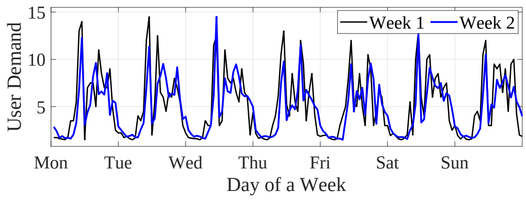
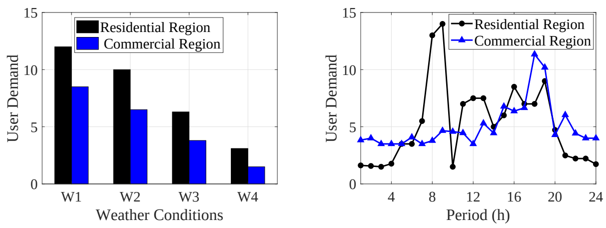


Fig. 7 Swapping demand over a two-week period.



(a) Demand under different weathers (b) Demand across different regions

Fig. 8 Swapping demand (a) under different weathers and (b) across different regions.

denoising process to generate samples that meet conditional constraints. Predicting demand distribution $d_{t,i} \sim q(d_{t,i})$ needs to model a probability density function $p_\theta(d_{t,i})$ that approximates $q(d_{t,i})$. Let k denotes the k-th diffusion step, where $k = 1, 2, \ldots, K$, and the forward diffusion process is defined by a Markov chain:

$$q(d_{t,i}^{1:K} | d_{t,i}^0) = \prod_{k=1}^{K} q(d_{t,i}^k | d_{t,i}^{k-1}), \tag{10}$$

where $q(d_{t,i}^k | d_{t,i}^{k-1}) = \mathcal{N}(\sqrt{1-\beta_k} d_{t,i}^{k-1}, \beta_k I)$ and $\beta_k \in (0,1)$ is the noise level. The demand distribution $q(d_{t,i}^k)$ at k-th step can be written in a close form as $q(d_{t,i}^k | d_{t,i}^0) = \mathcal{N}\left(d_{t,i}^k; \sqrt{\bar{\alpha}_k} d_{t,i}^0, (1-\bar{\alpha}_k)I\right)$, where $\alpha_k = 1 - \beta_k$ and $\bar{\alpha}_k = \prod_{k=1}^{K} \alpha_k$. The demand $d_{t,i}^k$ at k-th step can be described as $d_{t,i}^k = \sqrt{\bar{\alpha}_k} d_{t,i}^0 + \sqrt{1-\bar{\alpha}_k}\epsilon$, where ϵ sampled from a standard Gaussian noise $\epsilon \sim \mathcal{N}(0, I)$.

In the reverse denoising process, we can predict the swapping demand $d_{t,i}^0$ for BSS v_i at t through denoising Gaussian noise $d_{t,i}^K \sim \mathcal{N}(0, I)$ conditioned on spatiotemporal representation $h_{t,i}^{lg}$:

$$p_\theta(d_{t,i}^{0:K} | h_{t,i}^{lg}) = p(d_{t,i}^K) \prod_{k=1}^{K} p_\theta(d_{t,i}^{k-1} | d_{t,i}^k, h_{t,i}^{lg}), \tag{11}$$

$$p_\theta(d_{t,i}^{k-1} | d_{t,i}^k, h_{t,i}^{lg}) = \mathcal{N}(d_{t,i}^{k-1}; \mu_\theta(d_{t,i}^k, h_{t,i}^{lg}), \sigma_\theta(d_{t,i}^k, h_{t,i}^{lg}) I). \tag{12}$$

We train a denoising network ϵ_θ to predict the noise ϵ added to $d_{t,i}^0$ and we can obtain the μ_θ and σ_θ as below:

$$\mu_\theta(d_{t,i}^k, h_{t,i}^{lg}) = \frac{1}{\sqrt{\alpha_k}}(d_{t,i}^k - \frac{\beta_k}{1-\sqrt{\bar{\alpha}_k}} \epsilon_\theta(d_{t,i}^k, h_{t,i}^{lg})), \tag{13}$$

Algorithm 2: Sampling $d_{t,i}^0$

Input: $d_{t,i}^K \sim \mathcal{N}(0, I)$; $h_{t,i}^{lg}$;
for $k = K$ *to* 1 **do**
  $\tilde{\epsilon} \sim \mathcal{N}(0, I)$ if $k > 1$ else $\tilde{\epsilon} = 0$;
  $d_{t,i}^{k-1} = \frac{1}{\sqrt{\alpha_k}}(d_{t,i}^k - \frac{\beta_k}{1-\sqrt{\bar{\alpha}_k}} \epsilon_\theta(d_{t,i}^k, h_{t,i}^{lg})) + \sqrt{\beta_k}\tilde{\epsilon}$;
return $d_{t,i}^0$

$$\sigma_\theta(d_{t,i}^k, h_{t,i}^{lg}) = \beta_k \tilde{\epsilon}, \tag{14}$$

where $\tilde{\epsilon}$ is stochastic variables sampled from a standard Gaussian distribution. The denoising network ϵ_θ can be trained by the following loss function:

$$\mathcal{L}_1(\theta) = \mathbb{E}_{d_{t,i}^0, \epsilon, k}\left[\left\| \epsilon - \epsilon_\theta(\sqrt{\bar{\alpha}_k} d_{t,i}^0 + \sqrt{1-\bar{\alpha}_k}\epsilon, h_{t,i}^{lg}) \right\|^2\right]. \tag{15}$$

The pseudocode for the training procedure and the sampling procedure are presented in Alg. 1 and Alg. 2.

5.2 Deep reinforcement learning-based battery redistribution

As shown in Fig. 9, we propose a demand-oriented deep reinforcement learning model in this section. It captures environment features in battery swapping scenario through an Urban Knowledge Graph (UKG), describes the refinement of BSSs as MMDPs (see Sec. 5.2.1), and learns the optimal strategy through a multi-agent DRL (see Sec. 5.2.2).

5.2.1 Markov decision process

Formally, we model the BNR problem as MMDPs $\mathcal{G}$ for N BSS agents, which is defined as a quintuple $\mathcal{G} = (\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \beta)$, where $\mathcal{S}$ is the state space, $\mathcal{A}$ is the action space, $\mathcal{P}$ is the state transition function, $\mathcal{R}$ is the reward function, and β is the discount factor. The details of MDP $\mathcal{G}$ are shown below.

State $\mathcal{S}$. We observe the BSS network G, the swapping demand distribution D_t, multi-source environment features O_t, and the battery distribution N_t during time slot t. Then, we obtain the BSS agents' state space $s_t = \{s_{t,1}, s_{t,2}, \ldots, s_{t,I}\}$ through learning across stations and perceiving the surrounding environmental information.

Specifically, inspired by previous work[39,40], we construct a UKG to extract the complex environmental representations of the BSSs. The UKG G^K consists of an urban entity set E^K, a relation set R^K, and a fact set F^K, where $F^K = \{< s^k, r^k, t^k > | s^k, t^k \in E^K, r^k \in R^K\}$, denoting head entity s^k is connected with tail entity t^k via relation r^k. We capture the complex relationships among city entities, as shown in Table 1. Specifically, we model regions as entities in the

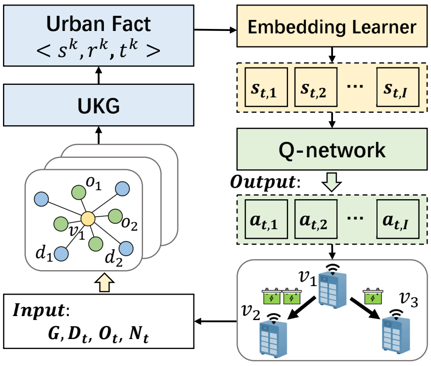


Fig. 9 The framework of demand-oriented deep reinforcement learning model.

UKG and describe the spatial adjacency relationships between regions using the *NeigWith* relation. Given that regions with similar functions tend to have similar swapping demand patterns, we connect these regions with the *SimilarFunc* relation. Environmental factors and their types are added as entities to the UKG, with the *CoverBy* and *TypeOf* relations representing the location and type of each factor, respectively. In addition, we introduce BSSs and batteries as entities into the UKG, and use the *LocateOn* and *CapaOf* relations to represent the location and capacity of each BSS. And we use *SatiBy* to describe that the swapping demand is satisfied by the BSS.

We employ a knowledge graph embedding model (*i.e.*, TuckER[41]) to learn entity representations that capture their semantic relationships. TuckER uses Tucker decomposition to evaluate the plausibility of a given fact, which is expressed as

$$\phi(s^k, r^k, t^k) = \mathcal{W} \times_1 \mathbf{s}^k \times_2 \mathbf{r}^k \times_3 \mathbf{t}^k, \tag{16}$$

where $\mathcal{W}$ is the core tensor to model the interaction between entities and relations. TuckER maximizes $\phi(s^k, r^k, t^k)$ to embed urban environmental features, from which we obtain the environmental embeddings for the BSS entity v_i at t, denoted as $s_{t,i}$.

Action $\mathcal{A}$. The action space of BSSs is $a_t = \{a_{t,1}, a_{t,2}, \ldots, a_{t,I}\}$, where each action $a_{t,i}$ represents agent BSS v_i transports $n_{t,i,j}$ batteries to BSS v_j during time slot t. We transform the action $a_{t,i}$ into a concatenation of an encoded vector $s_{t,j}$ and a scalar $n_{t,i,j}$, where $s_{t,j}$ is the target BSS and $n_{t,i,j}$ is the number of batteries transported. After performing the action, the number of batteries in BSS v_i is $n_{t-1,i} + \sum_{j=1}^{I} n_{t,j,i} - \sum_{j=1}^{I} n_{t,i,j}$. We avoid competition and conflicts between BSS agents by setting the battery dispatch and reception order. Specifically, we define a dispatch order for BSSs to send out batteries and a reception order for BSSs to receive batteries, which are dynamically adjusted based on the states of the BSSs. For each BSS v_i, we have: (i) To ensure cost efficiency and regional accessibility, BSS v_i needs to hold at least one battery and be limited to a maximum capacity range $n^m_{t,i}$. (ii) The number of batteries taken from v_i, denoted by $\sum_{j=1}^{I} n_{t,i,j}$, should not exceed the current number of batteries in v_i. (iii) Due to scheduling costs, we limit the total number of batteries scheduled across the BSS network to n^M. The state of BSS network at the $t+1$ follows transformation function $\mathcal{P}(s_{t+1}|s_t, a_t) : \mathcal{S} \times \mathcal{A} \to \mathcal{S}$.

Table 1 The details of relationships in UKG.

Relation	Entity Types	Information
NeigWith	(Region, Region)	Regions are neighbors
SimilarFunc	(Region, Region)	Functionally similar regions
CoverBy	(Env. factor, Region)	Env. factor in the region
TypeOf	(Env. factor, Type)	Type of Env. factor
LocateOn	(BSS, Region)	BSS in the region
CapaOf	(BSS, Battery)	Capacity of BSS
SatiBy	(Demand, BSS)	Demand for the BSS

Reward $\mathcal{R}$. Reward reflects the immediate sense of the action in a specific state, and is typically used to guide the agent's learning process. A common measurement is to evaluate the cumulative reward difference between taking an action and not. In the shared battery scheduling scenario, we define two types of rewards: (i) Reward of satisfying predicted user demand. (ii) The potential social reward generated by environmental impact. We assume each BSS prioritizes serving the nearest users, and users in a region always swap batteries at the local BSS when needed. Since the current battery scheduling affects the future state of the BSS network, it brings long-term positive rewards. After taking action $a_{t,i}$, the BSS agent v_i transport from state $s_{t,i}$ to state $s_{t+1,i}$, and its reward $r_{t,i}$ is shown in the following:

$$r_{t,i} = \min(d_{t,i}, n_{t-1,i} + \sum_{j=1}^{I} n_{t,j,i} - \sum_{j=1}^{I} n_{t,i,j}) + o_{t,i}. \tag{17}$$

Therefore, the reward of the BSS network is the sum of the rewards of BSSs, as follows:

$$r_t = \sum_{i=1}^{I} r_{t,i}. \tag{18}$$

State transition function $\mathcal{P}$. We map $\mathcal{P}$ to $\mathcal{S} \times \mathcal{A} \times \mathcal{S} \to [0, 1)$. $\mathcal{P}(s_{t+1}|s_t, a_t)$ indicates the transition probability for BSS network from the current state s_t to

Algorithm 3: Demand-oriented deep reinforcement learning algorithm

Input: Q; Q_w; λ; λ_ω; $\mathcal{B}$; $\mathcal{N}_t$;N_s;
for *episode* $= 1, 2, \ldots, C$ **do**
  Initialize state $s_{0,i}$;
  for *step* $t = T+1, T+2, \ldots, T+T'$ **do**
    Select and execute action $a_{t,i} = \arg\max_a Q_(s_{t,i}, a; \lambda) + \mathcal{N}_{t,i}$;
    $r_{t,i}, s_{t+1,i} \leftarrow a_{t,i}$;
    $\mathcal{B} \leftarrow (\{s_{t,i}, a_{t,i}, r_{t,i}, s_{t+1,i}\})$;
    Sample a random minibatch of X transitions $\{(s_x, a_x, r_x, s_{x+1})\}_{x=1,2,\ldots,X}$ from $\mathcal{B}$;
    Compute target value $Y_{t,i}$ according to Eq. (20);
    Update Q-network parameters λ by minimizing the loss according to Eq. (19);
    Update Target network parameters $\lambda_\omega \leftarrow \lambda$ every N_s steps;

the next state s_{t+1} based on actions a_t. Specifically, the BSS agents select actions a_t based on their current state s_t (including the BSS network, the swapping demand distribution, multi-source environment features, and the battery distribution), which alters the state to s_{t+1}. The BSS agents receive total reward r_{t+1} based on the state transition, allowing them to learn the effectiveness of different actions.

Discount factor β. The value of β determines how much the reinforcement learning agents focus on forward reward. β is usually taken in the range $[0, 1)$ to ensure that the final reward converges to a finite value. When $\beta = 0$, the agents focus only on immediate reward and ignore future reward.

5.2.2 DRL Implementation

The goal of BSS agents is to learn a policy that maximizes the expected cumulative future reward $\sum_{t=T+1}^{T+T'} \sum_{i=1}^{I} r_{t,i}$. We design a long-term value network to learn the optimal actions for BSS agents, *i.e.,* which action to take between two time periods. Specifically, we use the Q-network and Target network[42] to model the state transition from $s_{t,i}$ to $s_{t+1,i}$, with $Q\left(s_{t,i}, a_{t,i}; \lambda\right)$ and $Q_\omega\left(s_{t,i}, a_{t,i}; \lambda_\omega\right)$ representing the outputs, and λ and λ_ω as the network parameters. We randomly sample $(s_{t,i}, a_{t,i}, r_{t,i}, s_{t+1,i})$ from the experience replay $\mathcal{B}$, and the loss of Q-network is expressed as

$$\mathcal{L}_2(\lambda) = \mathbb{E}_{s_{t,i}, a_{t,i}, s_{t+1,i} \sim \mathcal{B}}[(Q(s_{t,i}, a_{t,i}; \lambda) - Y_{t,i})^2], \tag{19}$$

where $Y_{t,i}$ is the target value, which is expressed as

$$Y_{t,i} = r_{t,i} + \beta Q_\omega(s_{t+1,i}, \arg\max_a Q(s_{t+1,i}, a; \lambda); \lambda_\omega), \tag{20}$$

where $r_{t,i}$ is the reward of state transition. We set the learning rate $\eta \in [0, 1]$ and update parameters as follows:

$$Q(s_{t,i}, a_{t,i}; \lambda) = Q(s_{t,i}, a_{t,i}; \lambda) + \eta(Y_{t,i} - Q(s_{t,i}, a_{t,i}; \lambda)). \tag{21}$$

To ensure stable training, the parameters λ are copied to the Target network to update λ_ω every N_s step. To maintain exploration diversity and speed up convergence, we introduce a noise network that generates Gaussian noise with the same dimension as the action space, and add it to the action values. The training procedure for DRL agents is shown in Alg. 3.

6 Evaluation

We perform extensive experiments and answer the following three research questions:

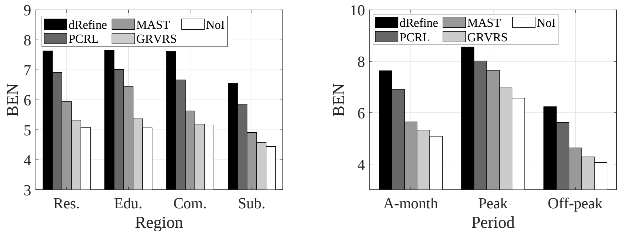
- **RQ1:** How does *dRefine* perform compared with existing baseline approaches?
- **RQ2:** How do the critical modules of *dRefine* impact the performance?
- **RQ3:** How do the key hyperparameters influence the performance of *dRefine*?

6.1 Experiment Settings

Dataset. We evaluate *dRefine* based on real-world datasets over six months (*i.e.,* from January 2024 to June 2024) from four major Chinese cities: Beijing, Shanghai, Chengdu, and Nanjing, which contain a total of 552,727 swapping records, generated by 41,358 batteries and 108,574 e-scooter users. We divide the cities into 2 km by 2 km urban regions. The detailed data format is given in Sec. 3.2.

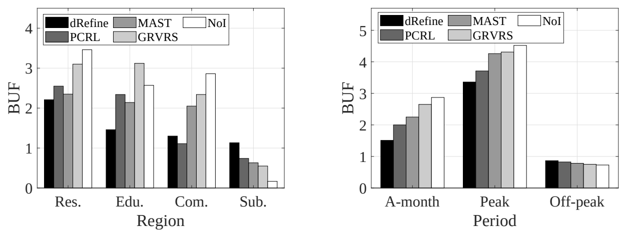
Baselines. We compare *dRefine* with the following baselines:

- *No Imbalance (NoI)*: It has been implemented across BSS networks and is currently in operation. Specifically, The average demand over the first seven days is utilized as the predicted demand, and batteries are redistributed accordingly.
- *GRVRS*[43]: A greedy-based algorithm designed for rebalancing shared resource distribution, achieving an effective balance between runtime efficiency and solution optimality.
- *MAST*[44]: A novel mathematical model based on the ant colony algorithm with soft time windows, specifically designed for resource allocation in shared systems. The model aims to minimize



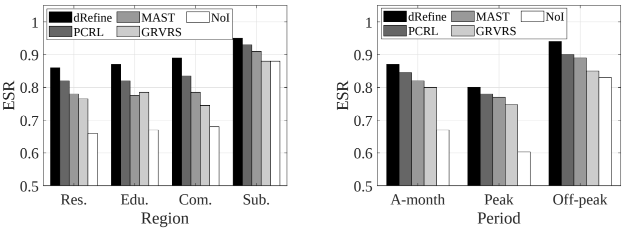
(a) BEN for different regions (b) BEN for different periods

Fig. 10 Performance evaluations of BEN in (a) different regions and (b) different periods.



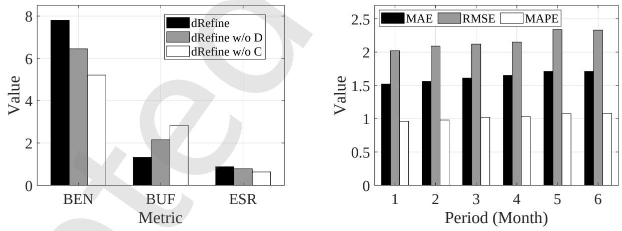
(a) BUF for different regions (b) BUF for different periods

Fig. 11 Performance evaluations of BUF in (a) different regions and (b) different periods.



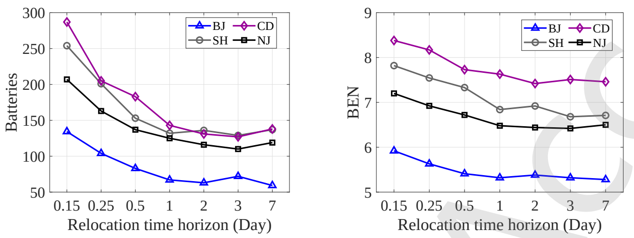
(a) ESR for different regions (b) ESR for different periods

Fig. 12 Performance evaluations of ESR in (a) different regions and (b) different periods.



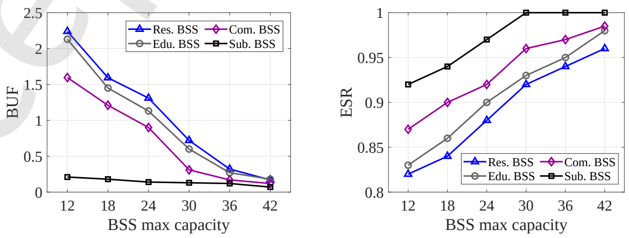
(a) Ablation study (b) Impact of prediction time horizon

Fig. 13 (a) Results for ablation study and (b) impact of prediction time horizon.



(a) Batteries for different relocation horizons (b) BEN for different relocation horizons

Fig. 14 Impact of relocation time horizon in terms of (a) relocated battery numbers and (b) BEN. (BJ: Beijing, SH: Shanghai, CD: Chengdu, NJ: Nanjing)



(a) BUF for different BSS max capacities (b) ESR for different BSS max capacities

Fig. 15 Impact of BSS max capacity in terms of (a) BUF and (d) ESR.

transportation costs while simultaneously reducing customer waiting times.

- *PCRL*[45]: A deep reinforcement learning approach designed to determine the optimal locations for stations and the ideal number of allocated resources.

Evaluation metrics. Since battery swapping takes only a few minutes, users' waiting time is negligible in most cases. Therefore, our evaluation primarily focuses on system performance and battery utilization. We use the following three metrics to evaluate *dRefine*:

- *Benefit (BEN)*: The BEN of BSSs is defined by Eq. 17 in Sec. 5.2.1, which incorporates the benefits of meeting predicted user demand and the potential social benefits. This metric is utilized to assess the operational efficiency of the BSS network following refinement.

- *Battery Usage Frequency (BUF)*: The BUF is defined as the ratio of user swapping demand $d_{t,i}$ to the number of batteries $n_{t,i}$ available at a BSS. This indicator is used to measure the utilization efficiency of battery resources.

- *Effective Swap Ratio (ESR)*: The ESR is defined as the ratio of the number of batteries removed with SoC above the threshold to the total number of batteries picked from the BSS $n_{t,i}$. The metric is key to evaluating the service quality of BSSs.

6.2 Experiment Results

Overall Performance. Figs. 10-12 show the result of our model and all baselines. Overall, *dRefine* achieves the best performance among all the baseline methods. Specifically, we make the following observations:

(1) *Performance on BEN*. As shown in Fig. 10a, we evaluate the BEN performance of algorithms across different functional regions, including

residential, educational, commercial, and suburban regions. The BEN of *dRefine* outperforms other algorithms by an average of 29.28%, demonstrating its ability to effectively consider critical factors in the BSS network and ensure successful refinement. The reinforcement learning-based method (PCRL) significantly outperforms other methods (*i.e.,* MAST, GRVRS, NoI), as it effectively addresses the high-dimensional state caused by large-scale agents in battery allocation process. We also analyze the BEN of three time periods after battery relocation: over a month, peak periods, and off-peak periods, and compare *dRefine* with other algorithms. Fig. 10b shows that *dRefine* outperforms the baselines. Notably, as swapping demand fluctuates regularly, the BEN is higher in peak periods and lower in off-peak periods.

(2) *Performance on BUF*. Fig. 11a shows the BUF of different algorithms for different regions, with values approaching 1 indicating efficient battery utilization. *dRefine* achieves the best performance, decreasing the BUF by an average of 31.22% compared to baseline methods, followed by PCRL. It shows that *dRefine* adapts to the environmental features of different regions well. MAST outperforms GRVRS because it considers soft time window constraints, while other rules remain the same as GRVRS, making BSS refinement more forward-looking. Additionally, we find that BUF is higher in regions with high user flow. In suburban regions with low user flow, particularly for NoI, BUF is significantly lower than 1, indicating that supply exceeds demand. From Fig. 11b, we find that BUF is higher during peak periods and lower during off-peak periods, demonstrating the importance of dynamically scheduling batteries.

(3) *Performance on ESR*. As shown in Fig. 12a, *dRefine* achieves the best performance in ESR, with an average improvement of 11.54%, effectively balancing supply and demand while providing more time for battery charging. Notably, in suburban regions, the ESR of the algorithms is significantly high due to long intervals between user visits and low demand. From a temporal perspective, Fig. 12b shows that the ESR of NoI is low during peak periods, and algorithms achieve noticeable improvements. In off-peak periods, the ESR of NoI is high, and algorithms achieve smaller improvements, especially in GRVRS.

Ablation Study. To further study the contribution of each module, we conduct an ablation study by removing key components of *dRefine*. The variants are: (i) *dRefine* w/o C, which replaces cross-region demand prediction model with the ConvLSTM. ConvLSTM[46] is a variant of LSTM that integrates convolutional layers, developed to tackle spatiotemporal sequence prediction tasks. (ii) *dRefine* w/o D, which substitutes demand-oriented deep reinforcement learning model with MAST. The results in Fig. 13a show that: (i) When removing prediction module, BEN decreases rapidly, and BUF increases. It shows that lacking accurate prediction of swapping demand leads to improper resource scheduling, which can worsen battery imbalance in urban regions and make it difficult to meet the demand promptly. (ii) *dRefine* w/o D outperforms *dRefine* w/o C because it still benefits from accurate demand prediction. However, without the ability to adaptively refine battery scheduling over time, it performs worse than *dRefine*. It indicates the importance of dynamically capturing and modeling the battery swapping environment in cities to achieve the long-term benefits of the BSS network.

Hyperparameter Analysis. We perform analyses of the hyperparameters of our model to understand their impact on performance.

(1) *Prediction horizon*. We assess the impact of the prediction horizon on the performance of the prediction module using standard metrics, *i.e.,* MAE, RMSE, and MAPE. As Fig. 13b shows, with the prediction time horizon increasing, the MAE, RMSE, and MAPE rise slowly, demonstrating the long-term effectiveness and robustness of the prediction module.

(2) *Relocation time horizon*. We evaluate the impact of relocation time horizon on model performance across four city datasets by extending it from 0.15 to 7 days. As depicted in Fig. 14a, the number of batteries to relocate decreases rapidly as relocation horizon increases, stabilizing once the horizon reaches 1 day. For instance, as the relocation horizon increases from 0.15 to 1, the number of batteries to relocate decreases by 46.94%, and from 1 to 7, it decreases by 4.97%. This is because user demand changes significantly between hours, while demand distribution across days is similar. In Fig. 14b, we find that as relocation horizon increases, BEN decreases and then stabilizes when the horizon reaches 1 day. This is because BEN depends on meeting user demand, which changes frequently. When batteries are not relocated frequently, some demands cannot be satisfied. In summary, while a higher scheduling frequency leads to greater scheduling costs, it also results in higher benefits.

(3) *Max capacity of BSSs*. We evaluate the impact of max capacity of BSSs on the performance of *dRefine* at different regions. In Fig. 15a, as the max capacity increases, the BUF first approaches 1 and then decreases towards 0. This indicates that a smaller capacity causes battery shortages, while excessive capacity leads to idle battery resources. Additionally, BUF is lower for BSSs in suburban regions due to lower user traffic, allowing for smaller-scale BSS construction. Fig. 15b shows that ESR at regions increases at different rates as the max capacity increases. This indicates that different regions require different BSS capacities, and battery swapping providers can adjust the set capacities of BSSs to balance costs and benefits.

7 Conclusion

In this paper, we explore the problem of BSS network refinement, *i.e.,* relocating batteries across stations to meet the fluctuating demands of e-scooters. To solve this problem, we propose *dRefine*, a novel data-driven battery swapping station network refinement system, which integrates two key modules: user demand prediction and battery relocation. For user demand prediction, we propose a spatiotemporal context representation-guided conditional diffusion model, using data from regions with established networks to predict demand in data-scarce regions. For battery relocation, we develop a demand-oriented deep reinforcement learning model that dynamically optimizes scheduling strategies to satisfy users' long-term demands. We conduct extensive experiments using real-world, large-scale battery swapping data, demonstrating that *dRefine* outperforms existing state-of-the-art methods.

Acknowledgement

This research is supported by the National Natural Science Foundation of China under Grant No. 62171454, 62441207, NSF China Projects No. 62172063, and NSF of Chongqing Projects No. CSTB2024NSCQ-JQX0020.

References

[1] H. Lu, Q. Zhang, and C. Xiang, "Battery swap station site selection algorithm based on deep reinforcement learning," in *Wireless Sensor Networks: 17th China Conference, CWSN 2023*, Dalian, China, 2023.

[2] "Electric scooter market size, share & industry analysis," 2024. [Online]. Available: https://www.fortunebusinessinsights.com/electric-scooter-market-102056.

[3] "Battery swapping market worth $11.8 billion by 2027," 2023. [Online]. Available: https://www.marketsandmarkets.com/PressReleases/battery-swapping.asp.

[4] G. Battapothula, C. Yammani, and S. Maheswarapu, "Multi-objective optimal scheduling of electric vehicle batteries in battery swapping station," in *2019 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, Bucharest, Romania, 2019, pp. 1–5.

[5] A. A. Shalaby, M. F. Shaaban, M. Mokhtar, H. H. Zeineldin, and E. F. El-Saadany, "A dynamic optimal battery swapping mechanism for electric vehicles using an LSTM-based rolling horizon approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 15 218–15 232, 2022.

[6] E. Zhou, Z. Li, D. Liu, C. Xiang, J. Chen, and W. Cheng, "Balancing electric scooter battery swapping network by spatio-temporal recommendation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 12, pp. 21 315–21 326, 2024.

[7] H. Wu, "A survey of battery swapping stations for electric vehicles: Operation modes and decision scenarios," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 10 163–10 185, 2022.

[8] C. Wang, J. Yang, N. Liu, and Y. Mao, "Study on siting and sizing of battery-switch station," in *2011 4th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT)*, Weihai, China, 2011, pp. 657–662.

[9] C. Xiang, W. Cheng, C. Lin, X. Zhang, D. Liu, X. Zheng, and Z. Li, "Lstaloc: A driver-oriented incentive mechanism for mobility-on-demand vehicular crowdsensing market," *IEEE Transactions on Mobile Computing*, vol. 23, no. 4, pp. 3106–3122, 2024.

[10] N. P. Mohapatra and S. R. Pattanaik, "Green internet of vehicles (giov): Applications, awareness, technologies and challenges," in *2023 IEEE 3rd International Conference on Sustainable Energy and Future Electric Transportation (SEFET)*, Bhubaneswar, India, 2023, pp. 01–07.

[11] P. Rani and R. Sharma, "Intelligent transportation system performance analysis of indoor and outdoor internet of vehicle (iov) applications towards 5g," *Tsinghua Science and Technology*, vol. 29, no. 6, pp. 1785–1795, 2024.

[12] G. Ramesh, A. K. Budati, S. Islam, L. A. Maghrabi, and A. Al-Atwai, "Artificial intelligence enabled future wireless electric vehicles with multi-model learning and decision making models," *Tsinghua Science and Technology*, vol. 29, no. 6, pp. 1776–1784, 2024.

[13] Z. Moghaddam, I. Ahmad, D. Habibi, and Q. V. Phung, "Smart charging strategy for electric vehicle charging stations," *IEEE Transactions on Transportation Electrification*, vol. 4, no. 1, pp. 76–88, 2017.

[14] G. Wang, Z. Fang, X. Xie, S. Wang, H. Sun, F. Zhang, Y. Liu, and D. Zhang, "Pricing-aware real-time charging scheduling and charging station expansion for large-scale electric buses," *ACM Transactions on Intelligent Systems and Technology*, vol. 12, no. 1, pp. 1–26, 2020.

[15] F. Ahmad, M. Saad Alam, I. Saad Alsaidan, and S. M. Shariff, "Battery swapping station for electric vehicles: opportunities and challenges," *IET Smart Grid*, vol. 3, no. 3, pp. 280–286, 2020.

[16] S. Wang, X. Zhu, G. Wang, Y. Liu, T. He, and D. Zhang, "eshare+: A data-driven balancing mechanism for bike sharing systems considering both quality of service and maintenance," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 10, pp. 10 497–10 513, 2023.

[17] J. Liu, L. Sun, W. Chen, and H. Xiong, "Rebalancing bike sharing systems: A multi-source data smart optimization," in *Proc. ACM SIGKDD*, Francisco, CA, USA, 2016, pp. 1005–1014.

[18] H. Wu, G. K. H. Pang, K. L. Choy, and H. Y. Lam, "An optimization model for electric vehicle battery charging at a battery swapping station," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 2, pp. 881–895, 2017.

[19] B. Du, Y. Tong, Z. Zhou, Q. Tao, and W. Zhou, "Demand-aware charger planning for electric vehicle sharing," in *Proc. ACM SIGKDD*, London, United Kingdom, 2018, pp. 1330–1338.

[20] Y.-L. Lv, S.-B. Xu, and J.-H. Xu, "A genetic algorithm based battery swapping stations localization model for electric transportation system," in *Proc. IEEE ICWAPR*, Adelaide, Australia, 2021, pp. 1–6.

[21] S. Xiaoyu, L. Wenyi, and L. Xiaolong, "Research on ordered charging of battery swapping station based on adaptive genetic algorithm," in *Proc. IEEE iSPEC*, Beijing, China, 2019, pp. 792–796.

[22] Y. Tao, M. Huang, and L. Yang, "Data-driven optimized layout of battery electric vehicle charging infrastructure," *Energy*, vol. 150, no. 1, pp. 735–744, 2018.

[23] Y. Zheng, Z. Y. Dong, Y. Xu, K. Meng, J. H. Zhao, and J. Qiu, "Electric vehicle battery charging/swap stations in distribution systems: Comparison study and optimal planning," *IEEE Transactions on Power Systems*, vol. 29, no. 1, pp. 221–229, 2014.

[24] S. Chirag Anirudh, A. Bagrodia, A. S. Ubaradka, and V. K. Tumuluru, "E-bike station management using deep reinforcement learning," in *2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT)*, Vellore, India, 2024, pp. 1–6.

[25] C. Wang, Y. Song, Y. Wei, G. Fan, H. Jin, and F. Zhang, "Towards minimum fleet for ridesharing-aware mobility-on-demand systems," in *Proc. IEEE INFOCOM*, Vancouver, BC, Canada, 2021, pp. 1–10.

[26] S. Mishra, A. Mondal, and S. Mondal, "Charging station siting and sizing considering uncertainty in electric vehicle charging demand distribution,"

IEEE Transactions on Intelligent Transportation Systems, pp. 1–13, 2024 (Early Access).

[27] G. Koolman, M. Stecca, and P. Bauer, "Optimal battery energy storage system sizing for demand charge management in ev fast charging stations," in *Proc. IEEE ITEC*, Chicago, IL, USA, 2021, pp. 588–594.

[28] L. Zhao, S. Shen, and Z. Zhao, "Planning decentralized battery-swapping recharging facilities for e-bike sharing systems," *Sustainable Cities and Society*, vol. 101, p. 105118, 2024.

[29] C. He, J. Zhu, S. Li, Z. Chen, and W. Wu, "Sizing and locating planning of ev centralized-battery-charging-station considering battery logistics system," in *Proc. IEEE CIEEC*, Wuhan, China, 2021, pp. 1–6.

[30] T.-Y. Zhang, Y. Yang, Y.-T. Zhu, E.-J. Yao, and K.-Q. Wu, "Deploying public charging stations for battery electric vehicles on the expressway network based on dynamic charging demand," *IEEE Transactions on Transportation Electrification*, vol. 8, no. 2, pp. 2531–2548, 2022.

[31] M. R. Sarker, H. Pandžić, and M. A. Ortega-Vazquez, "Optimal operation and services scheduling for an electric vehicle battery swapping station," *IEEE Transactions on Power Systems*, vol. 30, no. 2, pp. 901–910, 2015.

[32] S. Wang, C. Shao, C. Zhuge, M. Sun, P. Wang, and X. Yang, "Deploying battery swap stations for electric freight vehicles based on trajectory data analysis," *IEEE Transactions on Transportation Electrification*, vol. 8, no. 3, pp. 3782–3800, 2022.

[33] "Sichuan xianglilai technology co., ltd." 2023. [Online]. Available: https://www.scxll.cn/.

[34] M. L. Puterman, "Markov decision processes," *Handbooks in Operations Research and Management Science*, vol. 2, pp. 331–434, 1990.

[35] W. Cheng, H. Lu, C. Xiang, D. Liu, and T. Xiang, "Breaking 'chicken-egg': Cross-city battery swap demand prediction via knowledge-guided diffusion," in *Proc. IEEE INFOCOM*, London, United Kingdom, 2025, pp. 1–10.

[36] Y. Liu, Z. Zhou, Y. Li, and D. Jin, "Urban knowledge graph aided mobile user profiling," *ACM Transactions on Knowledge Discovery from Data*, vol. 18, no. 1, pp. 1–30, 2024.

[37] H. Wang, Q. Yu, Y. Liu, D. Jin, and Y. Li, "Spatio-temporal urban knowledge graph enabled mobility prediction," in *Proc. ACM UbiComp*, New York, NY, USA, 2021, pp. 1–24.

[38] S. Hui, H. Wang, T. Li, X. Yang, X. Wang, J. Feng, L. Zhu, C. Deng, P. Hui, D. Jin, and Y. Li, "Large-scale urban cellular traffic generation via knowledge-enhanced GANs with multi-periodic patterns," in *Proc. ACM KDD*, Long Beach, CA, USA, 2023, pp. 4195–4206.

[39] Z. Zhou, J. Ding, Y. Liu, D. Jin, and Y. Li, "Towards generative modeling of urban flow through knowledge-enhanced denoising diffusion," in *Proc. ACM SIGSPATIAL*, Hamburg, Germany, 2023, pp. 1–12.

[40] S. Zhang, T. Li, S. Hui, G. Li, Y. Liang, L. Yu, D. Jin, and Y. Li, "Deep transfer learning for city-scale cellular traffic generation through urban knowledge graph," in *Proc. ACM SIGKDD*, Long Beach, CA, USA, 2023, pp. 4842–4851.

[41] I. Balazevic, C. Allen, and T. Hospedales, "Tucker: Tensor factorization for knowledge graph completion," in *Proc. EMNLP-IJCNLP*, Hong Kong, China, 2019, pp. 1–11.

[42] H. v. Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double q-learning," in *Proc. ACM AAAI*, Phoenix, AZ, USA, 2016, pp. 2094–2100.

[43] Y. Duan, J. Wu, and H. Zheng, "A greedy approach for vehicle routing when rebalancing bike sharing systems," in *2018 IEEE Global Communications Conference (GLOBECOM)*, Abu Dhabi, United Arab Emirates, 2018, pp. 1–7.

[44] D. Shenshen, C. Yun, and X. Hegen, "Research on scheduling optimization of free-floating bike sharing based on improved ant colony algorithm," in *2020 3rd World Conference on Mechanical Engineering and Intelligent Manufacturing (WCMEIM)*, Shanghai, China, 2020, pp. 699–702.

[45] L. von Wahl, N. Tempelmeier, A. Sao, and E. Demidova, "Reinforcement learning-based placement of charging stations in urban road networks," in *Proc. ACM SIGKDD*, Washington, DC, USA, 2022, pp. 3992–4000.

[46] Z. Lin, M. Li, Z. Zheng, Y. Cheng, and C. Yuan, "Self-attention convLSTM for spatiotemporal prediction," in *Proc. AAAI*, New York, NY, USA, 2020, pp. 11 531–11 538.

Han Lu received the B.E. degree in IoT and engineering from China University of Petroleum (East China), Qingdao, China, in 2023. She is currently pursing the M.S. degree in Chongqing University. Her current research interests include urban computing and IoT.

Wenhui Cheng received the B.E. degree in Mechanical and Electronic Engineering from East China Jiaotong University, China, in 2020. He is currently pursuing a Ph.D. degree at Chongqing University. His research interests include generative AI, machine learning, and urban computing.

Ruixue Huang received the B.E. degree in computer science and engineering from Chongqing University, Chongqing, China, in 2022. She is currently pursing the M.S. degree in Chongqing University. Her current research interests include vehicular crowdsensing and IoT.

Chaocan Xiang is a full professor at the College of Computer Science, Chongqing University, Chongqing, China. He received the BS and Ph.D. degrees in computer science and engineering from the Nanjing Institute of Communication Engineering, China, in 2009 and 2014, respectively. He studied at the Real-Time Computing Lab, the University of Michigan, Ann Arbor, MI, USA in 2017. His current research interests include crowd-sensing networks and IoT. He has published more than 30 research papers in important conferences and journals, such as ACM UbiComp, IEEE INFOCOM, IEEE/ACM TON, IEEE TMC, IEEE TPDS, IEEE T-ITS, and ACM TOSN.

Xiao Zheng is currently a Professor with the School of Computer Science and Technology, Anhui University of Technology, China. He received a B.S. degree from Anhui University, China, in 1997, an M.S. degree from Zhejiang University of Science and Technology, China, in 2003, and a Ph.D. degree in computer science and technology from Southeast University, China, in 2014.

Qing Zhou is a full professor at the College of Computer Science, Chongqing University, Chongqing, China. He got his PhD in Computer Science from Chongqing University in 2008. His current research interests include computational intelligence, data mining and educational technology.

Jianchao Zheng received the B.S. degree in communications engineering and the Ph.D. degree in communications and information systems from the College of Communications Engineering, PLA University of Science and Technology, Nanjing, China, in 2010 and 2016, respectively. From 2015 to 2016, he was a Visiting Scholar with the Broad-band Communications Research Group, Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. He is currently an Associate Research Fellow with the Academy of Military Sciences, Beijing, China. He has authored or coauthored several papers in international conferences and reputed journals in his research area. His research interests focus on communications and computing networks, game theory, learning theory, and optimization techniques.