IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 20 January 2023, accepted 28 January 2023, date of publication 2 February 2023, date of current version 10 February 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3241799

RESEARCH ARTICLE

An Intelligent Rebalance System for Tidal Phenomenon of Dockless Bicycle-Sharing

LYUCHAO LIAO[1], (Senior Member, IEEE), BEN LI[1], DEJUAN HUANG[1], ZHU XIAO[2,3], (Senior Member, IEEE), AND QI ZHENG[1]

[1]School of Transportation, Fujian University of Technology, Fuzhou 350118, China
[2]College of Computer Science and Electronic Engineering, Hunan University, Changsha 410082, China
[3]Shenzhen Research Institute, Hunan University, Shenzhen 518055, China

Corresponding authors: Ben Li (liben0210@outlook.com) and Zhu Xiao (zhxiao@hnu.edu.cn)

This work was supported in part by the Projects of the National Natural Science Foundation of China under Grant 41971340 and Grant 62272152; in part by the Fujian Provincial Department of Science and Technology under Grant 2021Y4019, Grant 2020D002, Grant 2020L3014, and Grant 2019I0019; in part by the Fujian Provincial Universities Engineering Research Center for Intelligent Driving Technology (FJUT) under Grant KF-J21012; in part by the Key Research and Development Project of Hunan Province of China under Grant 2022GK2020; and in part by the Shenzhen Science and Technology Program under Grant JCYJ20220530160408019.

ABSTRACT With the advantages of flexible parking locations and convenient cycling, Dockless Bicycle-sharing (DBS) has become increasingly popular worldwide. However, with the massive increase of DBSs and electric fences, DBS systems face several challenges: (1) the hardness of identifying the DBS tidal zones; (2) the difficulty of accurately evaluating and identifying overload fences; (3) the issues of rebalancing DBS in time. To deal with these challenges, we propose a Dockless Bicycle-sharing Dynamic Rebalance (DBSDR) system to dynamically provide the optimal bicycle guidance for the DBS network. The DBSDR system contains three modules: DBS tidal zone identification, evaluation framework of overload fences, and DBS dynamic guidance. For DBS tidal zone identification, tidal zone identification and location from each fence with bicycle flows are provided with the HDBSCAN clustering method. The evaluation framework, covering DBS flows and the parking demand density, is proposed to assess the characteristics of overload fences. Finally, a DBS dynamic guidance method is provided to balance DBS for the tidal phenomenon with guiding users to the optimal target fence. Extensive experiments conducted on real-world DBS datasets show the effectiveness and accuracy of rebalancing the tidal phenomenon in the DBS system.

INDEX TERMS Dockless Bicycle-sharing system, bicycle rebalancing, HDBSCAN, user-based guidance.

I. INTRODUCTION

In recent years, energy waste and environmental pollution have become more and more serious, and greenhouse gas emissions are one of the important factors [1], [2]. The rapid growth of bicycle-sharing provides an excellent opportunity to meet the needs of urban residents for green travel [3], [4]. In fact, for the public, the first/last mile problem is often the most immediate, and bicycle-sharing has the advantages of zero carbon emissions and convenient cycling [5], [6], [7]. Therefore, bicycle-sharing has obvious advantages in short-distance travel and is widely used as public transportation to solve the first/last mile problem [8], [9].

The associate editor coordinating the review of this manuscript and approving it for publication was Shaohua Wan.

At present, there are two main bicycle-sharing systems, station/dock-based bicycle-sharing (SDBS) system [10], [11] and Dockless Bicycle-sharing (DBS) system [12], [13], [14]. In the SDBS system, each station has multiple fixed bicycle docks, which greatly limits the movement of the station and the increase of bicycles, and prevents users from renting and parking bicycles [15]. The DBS system deploys bicycle-sharing at flexible electric fences rather than fixed stations where users can park their bicycles anywhere near the fences [16]. However, numerous bicycles are parked on some fences, while others have few bicycles, resulting in the tidal phenomenon (no bicycles' renting and parking in the morning and evening peak hours). Meanwhile, massive overload fences and idle fences in tidal zones lead to unbalance between supply and demand. Therefore, a reasonable guidance method is crucial for DBS.

Abound efforts have been devoted to the DBS [17], [18], [19], such as bicycle demand prediction, station layout modeling, and bicycle rebalance. However, balancing bicycle-sharing distribution in real-time is still challenging. Bicycle demand prediction is often used to predict bicycle flow during a period to guide staff to redistribute bicycles from overload fences to idle fences, which is hard to be efficient. Moreover, there is no clear evaluation framework for overload fences, making it hard to determine the specific overload fences accurately [5], [6], [7], [8], [9], [10], [11], [12], [13], [14]. Furthermore, based on the worker scheduling model, which cannot meet the needs of people's parking in time, this will lead to excessive operating costs. Significantly, the tidal zones are not well identified to evaluate overload fences further. However, the identification of overload fences can well solve the tidal phenomenon.

In this work, we propose a DBSDR system to guide users to park adjacent idle fences for rebalancing DBS. Unlike traditional methods based on a worker scheduling system, the DBSDR system is proposed to dynamically identify the tidal zones, evaluate the overload fences and rebalance them with a recommendation mechanism. More significantly, the system can provide users with the most suitable parking fences at any time from five target fence indicators (DBS flows of the recommended parking fences, density, active level, distances, and whether to cross the road) when they need to park. Meanwhile, we also took into account the global balance, guiding the bicycles of the overload fences to the adjacent idle fences to avoid excessive bicycles to park the same empty fence.

The main contributions of this work are summarized as follows:

- A method to identify the tidal zones with HDBSCAN is proposed, which meets the characteristics of the DBS dataset and combines the time-varying and spatial continuity.
- An evaluation framework is provided to assess the characteristics of overload fences and take it as the basis for identification.
- We propose a DBSDR system, which provides users with five target fence indicators to guide users in balancing the bicycle system for the tidal phenomenon. Extensive experiments on Xiamen's DBS dataset show the effectiveness of the DBSDR system.

The rest of the paper is organized as follows: In Section II, we summarize the related research on bicycle-sharing scheduling in recent years; notation and definition are provided in Section III. Then, we describe the methodology framework and the detailed three tasks in Section IV, and we conduct extensive experiments and discuss the experimental results in Section V. Finally, this work is concluded in Section VI.

II. RELATED WORKS

With the rapid development of the bicycle-sharing system in the last two decades, the DBS system particularly prospered in the past five years in China, and a large number of studies on DBS have emerged. Increasing attention has been paid to bicycle-sharing demand prediction and DBS rebalance.

A. BICYCLE-SHARING DEMAND PREDICTION

Accurate estimations for bicycle-sharing parking demands are crucial to managing and rebalancing bicycles [18], [20]. In general, the bicycle-sharing demand prediction can be conducted in two steps: the spatial-temporal zone division [21] and prediction algorithm design. For example, Chen et al. [22] built a weighted correlation network to support the application of geographically-constrained clustering for overdemand cluster prediction. Similarly, Huang et al. [23] further proposed a Two-Stage Station Clustering algorithm to cluster the central stations and common stations before predicting. In summary, the identification of the spatial-temporal zones is the basis of demand prediction. However, they lack an accurate evaluation framework for overload fences, and cannot accurately identify overload fences.

Numerous machine learning [24], [25] and deep learning methods [26], [27], [28], [29], [30] are applied to capture the spatial-temporal dependence of bicycle-sharing demand. For example, Gated Graph Neural Network (GGNN) was introduced to dynamically predict the bicycle station layout, the number of bicycles, and bicycle dispatching in work [31]. Although demand prediction algorithms can help predict the demand for electric fences, an inappropriate prediction method may mislead the user's judgment and often lead to locally optimal, guiding massive users to the same empty fence and resulting in a new overload fence.

B. BICYCLE-SHARING REBALANCE

Bicycle-sharing rebalance refers to the process in which the bicycles are relocated from the overload fences to the idle ones. A bicycle rebalances problem (BRP) is crucial to the rebalance operation of the system. Therefore, abundant studies have focused on static BRP (SBRP), which means repositioning bicycles at a fixed period [32], [33]. For example, Liu et al. [34] proposed an Adaptive Capacity Constrained K-centers Clustering (AdaCCKC) algorithm to divide outlier stations, reducing the largescale bicycle routing problems to the inner cluster one bicycle routing problem. Furthermore, swarm intelligence algorithms [35], [36] are also used to optimize the bicycle-sharing rebalancing process.

Dynamic BRP (DBRP) is widely studied and discussed to optimize the bicycle-sharing rebalance of the system. DBRP means the whole system is constantly updated, and the rebalancing schemes can be continuously adjusted [37], [38]. For example, Zhang et al. [37] proposed a zone-based two-stage rebalancing model, regulating zones with sufficient and deficient bikes two-stage. Tian et al. [39] designed a new framework to solve DBRP, which contains two sections: dynamic rebalancing of the inner station and static rebalancing among stations. Recently, user incentives-based bicycle-sharing rebalance was adopted to determine the optimal incentive scheme through reinforcement learning [12], [40].

However, these studies did not well identify the tidal zones and evaluate overload fences. The guidance on the overload fences and parking management scheme remains unclear. Therefore, we propose a DBSDR system to analyze the tidal zones and overload fences from the electric fence data and the DBS dataset. Recommendations with five target fence indicators are provided for parking guidance based on the overload fence identification results. Then, the balance of DSB system is achieved dynamically.

III. PRELIMINARIES

In this section, we first introduce the notations and definitions. Then we briefly introduce the geohash encoding.

A. NOTATION AND DEFINITION

For ease of illustration, we first summarize the notations and definitions.

TABLE 1. List of notations.

Symbols	Interpretation
M, N	The number of fences and existing bicycles of each fence.
C	The number of clusters of the fences.
G	The number of fences in the corresponding geohash encoding.
S	The size of each fence.

TABLE 2. List of indicators.

Indicators	Name of indicators	Description
F_{out}	Departure flow	The total number of bicycles outflow in this fence.
F_{in}	Arrival flow	The total number of bicycles inflow in this fence.
P_m	Remain flow	The remaining total number of bicycles in this fence.
A_m	Active level	The total number of active days in this fence.
L_m	Guidance distance	The distance between the bicycle and the recommended parking fences.
D_m	Demand density	The total number of bicycles per unit area.

Definition 1: Tidal phenomenon. The tidal phenomenon is defined as the phenomenon that bike-sharing cannot be rented or parked in some areas during morning and evening rush hours.

Definition 2: The active level of the fence. It describes the active days of $P_m > 0$ for the fence. The different active levels represent the degree to of demand exceeds supply.

B. GEOHASH ENCODING

Geohash (Geographical hashing) encoding is an algorithm that converts two-dimensional latitude and longitude data into a string, which is one of the most widely used urban address encodings. Geohash represents equal-length and equal-width squares. The longer the string length, the higher the accuracy.

$$Geocode = geohash\left(X_{lat}, X_{lon}\right) \tag{1}$$

In this work, a Geohash represents an electric fence or a bicycle location. For example, $Geocode = geohash$ $(24.468531, 118.098985) = ws7gpqm$. Users could publish address codes to protect privacy, providing their location without exposing precise coordinates. The encoding length needs to be selected according to the data. In this work, a spatial index based on Geohash could also help to improve the extraction efficiency of spatial data.

IV. METHODOLOGIES

In this section, we attempt to deal with three tasks: How are tidal zones distributed in space? How to evaluate and identify overload fences? How to guide users to park Bicycles for "peak clipping and valley filling"?

A. METHODOLOGY FRAMEWORK

The methodology framework is shown in Figure 1. The method collects multisource data, including DBS trip data, electric fence data, and DBS order data, to support DBS tidal zone identification, overload fences identification, and the guidance system.

1) TASK 1: DBS TIDAL ZONES IDENTIFICATION

Each trip's origin (O) and destination (D) are extracted from DBS order data. The longitude and latitude of each fence center are packed as the input of HDBSCAN. The determination of neighbor parameters of HDBSCAN is obtained by the KNN method. Finally, DBS tidal zones are identified.

2) TASK 2: EVALUATION FRAMEWORK OF OVERLOAD FENCES

The overload fences (OFs) are the critical areas for rebalancing the DBS. Meanwhile, the government also pays great attention to the bicycle parking demand in OFs. Therefore, the OFs obtained from Task 1 DBS tidal zones are chosen as Task 2's main rebalancing areas. Two indicators are extracted from the collected data to measure the OFs with high DBS flows and high parking demand density. For uniformity, DBS tidal zones are assessed similarly to overload fences. The identification result is used to guide parking management schemes.

3) TASK 3: DBS DYNAMIC GUIDANCE

Based on the rebalancing areas identified by the above two tasks, we propose a method to solve the rebalancing problem based on guiding users by five indicators. It can help "peak clipping and valley filling" for the overload fences while liberating a large amount of labor and saving scheduling costs. Furthermore, we took extensive factors into account for achieving global balance and preventing local optimums.

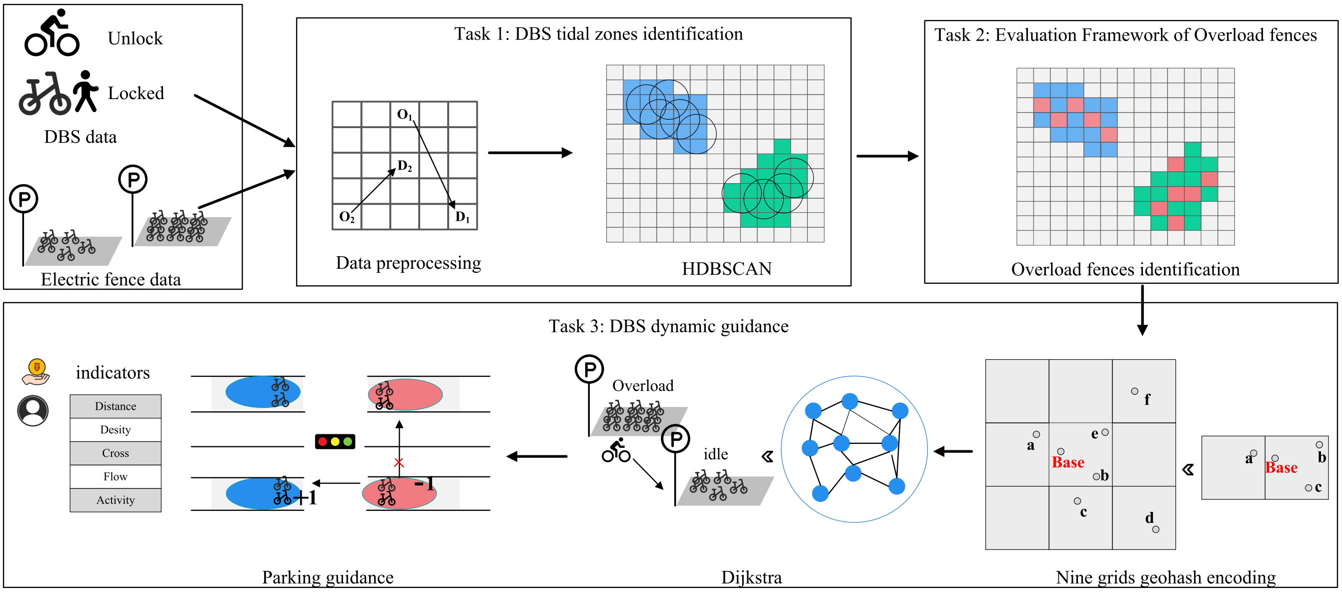


FIGURE 1. **The framework of DBSDR. Task 1: DBS tidal zones identification. Each trip's origin (O) and destination (D) are extracted from DBS order data. The longitude and latitude of each fence center are packed as the input of HDBSCAN to obtain the DBS tidal zones. Task 2: Evaluation Framework of Overload Fences. Two indicators are extracted from the collected data to measure the OFs with high DBS flows and high parking demand density. Task 3: DBS dynamic guidance. We extract the neighbor fences from the nine grids near each bicycle. The Dijkstra algorithm is employed to calculate the shortest distance between the shared bicycle and these fences in the directed weighted graph. Finally, user obtains the nearest fence based on guiding users by five indicators.**

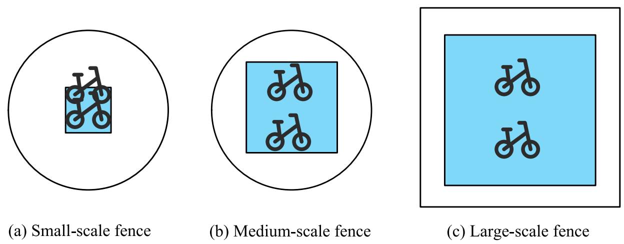


FIGURE 2. **The division of three different fence tolerance ranges.**

B. IDENTIFICATION OF DBS TIDAL ZONES

A DBS tidal zone is where many shared bicycles flow in or out in a specific time range. The main challenges faced by DBS operators and regulators are the DBS tidal zone identification. We propose an HDBSCAN-based method to identify the DBS tidal zones in this work.

Firstly, we set a suitable geohash encoding length to match the appropriate fence areas. The bicycles that are too far from the fences (more than 306 meters) will be taken as outliers and ongoing situations. These data can be regarded as part of the "disorderly parking" data in real world. Different fences have different areas, and their tolerance ranges could be different. Tolerance ranges of large fences are much more difficult than that of small and medium-sized fences, and special treatment of large fences will reduce the tolerable parking distance to make the results more accurate. Figure 2 shows the division of three different fence tolerance ranges. Based on three different fence tolerance ranges, we calculate the distance between the bicycles and the center of the fence and divide the order data into valid and invalid orders. Invalid orders are defined as points 40 meters from the parking fence center (for small and medium-sized fences) or 20 meters from the parking boundary (for large fences).

The tidal zone owns the characteristics of multiple domain-density maximums (MDDM), varying density distribution (VDD), and equilibrium distribution (ED). Specifically, bicycle-sharing has a different distribution in different periods, with time-varying. It is in line with the VDD feature. In addition, inflow and outflow in the stable distribution region will reach the balance between supply and demand, belonging to ED. Moreover, the tidal zones may have multiple domain density maximums during the morning peak. The HDBSCAN algorithm can obtain more reasonable clustering results for data with MDDM, ED, and VDD features. Therefore, we employ HDBSCAN to cluster the tidal zones.

HDBSCAN, based on density clustering combined with hierarchical analysis largely meets data features, whose input parameters are min_cluster_size and min_samples [41]. Specifically, there are some key definitions in the HDBSCAN, which are as follows: (1) Core distance: the distance between the sample point and the K-th nearest sample point; (2) mutual reachability distance: the value is the maximum value of the core distance of two sample points and the distance between two sample points. The mutual reachability distance can be obtained with Equation (2):

$$d_{mr}\left(a,b\right)=max\left\{d_c\left(a\right),d_c\left(b\right),d\left(a,b\right)\right\}, \qquad (2)$$

where $d\left(a,b\right)$ is the Euclidean distance between point a and point b, which increases the adaptability and robustness for different density regions.

To cluster the tidal zones, we first establish a minimum spanning tree with the mutual reachable distance between sample points as the edge and transform the tree into a hierarchical structure. Then, the input parameter min_cluster_size is employed to generate the compressed cluster tree. Finally, the density-adaptive clustering results are generated through a stability function.

Let φ_i and ψ_i be the latitude and longitude of the fence ν_i; we can use the Haversine method to calculate the distance between fences $\nu_i\,(\varphi_i, \psi_i)$ and $\nu_j\,(\varphi_j, \psi_j)$:

$$d_{ij} = 2R \cdot \sin\sqrt{H\left(\frac{d}{R}\right)} = 2R \cdot \times \sin\sqrt{H\left(\left|\psi_i - \psi_j\right|\right) + \cos\left(\psi_i\right)\cos\left(\psi_j\right) H\left(\left|\varphi_i - \varphi_j\right|\right)} \tag{3}$$

where R is the radius of the earth, usually set to 6371.0 km, and the Haversine function $H(\alpha)$ is defined as:

$$H(\alpha) = sin^2\left(\frac{\alpha}{2}\right) = \frac{1}{2}\left(1 - cos(\alpha)\right). \tag{4}$$

C. EVALUATION FRAMEWORK OF OVERLOAD FENCES

In this section, we develop an evaluation framework covering DBS flows P_m and the parking demand density D_m to assess the characteristics of overload fences, as shown in Table 2.

DBS flows of the fences represent the level of demand in a place and reflect the state of a fence. In general, it is divided into positive flows and negative flows (i.e., $P_m > 0$ and $P_m < 0$). The former means that the supply and demand can achieve a dynamic balance in the fence, while the latter means that staff is required to dispatch bicycles from other fences.

The parking demand density is also essential to reflect the congestion state of the fence. For example, the fence area with the same flow could be very different, in which the larger fence area will lead to a lower density.

1) DBS FLOW

DBS flow is an important index for OFs evaluation. Compared with docked bicycle-sharing, the parking of dockless bicycle-sharing tends to be more dispersed. DBS parking has become more restricted since implementing electric fences in the real world. The electric fences that specify the area that bicycles can be returned to and can be regarded as the parking supply location.

DBS flow is not only the basis for DBS tidal zones identification but also a key evaluation index of overload fences. Although tidal zones represent areas with high demand, the demand among these tidal zones is still uneven. The DBS flow is calculated as follows. As an example, a record of OD data will generate a departure demand at the origin and an arrival demand at the destination. The parking DBS flows P_m of this fence at time t could be calculated as follows:

$$P_m = N + F_{in} - F_{out}, \tag{5}$$

where P_m denotes the parking demand; F_{in}, F_{out} represent DBS trips arriving in electric fences and trips departing from electric fences, respectively. N is the number of existing bicycles of electric fences.

2) THE PARKING DEMAND DENSITY

The spatial-temporal variation of DBS demand density generally leads to the different congestion states of the fence. Demand density is defined as the ratio between the parking demand and the fenced area. Therefore, the parking demand density D_m can be calculated as follows:

$$S_m = length_m \cdot width_m \tag{6}$$

$$D_m = \frac{P_m}{S_m}, \tag{7}$$

where $length_m$ and $width_m$ represent the length and width of the M-th fence, respectively; S_m is the area of the M-th fence, and D_m is the bicycle density of the M-th fence.

To identify the overload fences accurately, we sum up the Z-score normalized the parking demand density and the parking demand to get the final overload fences. The comprehensive index O_m can be calculated as follows:

$$zscore(\cdot) = \frac{x - \mu}{\sigma} \tag{8}$$

$$O_m = zscore(P_m) + zscore(D_m), \tag{9}$$

where $zscore(\cdot)$ is the standardized equation; O_m is the comprehensive index for the M-th fence.

Evaluation and identification of overload fences using evaluation framework after identification of DBS tidal zones. Algorithm 1 describes the detailed steps of identification of overload fences.

D. DBS DYNAMIC GUIDANCE

We can obtain the location of overload fences based on the identification results of overload fences. As mentioned above, users could park their bicycles anywhere near their destination in a traditional DBS network. Generally, users' disorder parking will result in overload fences. Therefore, we need to guide users to park bicycles for "peak clipping and valley filling".

Firstly, we use geohash to encode the bicycles in overload fences and all the electric fences and match them with nine grids, respectively. We extract the neighbor fences from the nine grids near each bicycle. Then, remove or refuse bicycle parking without neighbor fences to prevent the user from parking disorderly. Secondly, with each fence as a node and the distance between them as weight, it can be seen as a directed weighted graph. And then, the Dijkstra algorithm is employed to calculate the shortest distance between the shared bicycle and these fences in the directed weighted graph. Finally, DBS dynamic guidance system obtains the nearest fence hash encoding.

In the case of the isolation belt, we match bicycles with the road network to determine whether it is necessary to cross the road to reach the nearest fence and provide it to the user. It is

Algorithm 1 Identification of Overload Fences

Input:

X : The bicycle dataset and the electric fence dataset for clustering;

θ : The threshold for the comprehensive index of each fence;

Output:

The overload fences.

```
extract all bicycle locations from X;
filter outers and obtain bicycle locations and fences locations;
the clusters C of the fences by HDBSCAN;
for each cluster c_i in clusters C do
    for each fence in the cluster c_i do
        calculate N, F_in, F_out of each fence for the BS flows P_m;
        calculate S_m of each fence for the parking demand density D_m;
        calculate the comprehensive index O_m of each fence using Equation (9);
        if O_m > θ then
            obtain the overload fences;
        end if
    end for
end for
return overload fences.
```

not only convenient for people, but also ensures the safety of users. Furthermore, the active level of the fence is also crucial, which describes the active days of $P_m > 0$ for the fence, as shown in Equation (10). The different active levels represent the degree of demand exceeding supply. For example, after the original number of bicycles in different fences flows out, the staff manually dispatches bicycles from other fences.

$$A_m = \sum_{t=1}^{T} a_m \begin{cases} 1 & if\ P_m > 0 \\ 0 & otherwise \end{cases} \tag{10}$$

where T is the total sampling days, and a_m is the active days for the M-th fence.

Users could also set their acceptable range, generally about 200 - 500 meters. DBS dynamic guidance system sums up standardized the comprehensive index weight of the electric fence, active level, and the guidance distance to obtain the recommended index, as shown in Equation (11). The system sorts the recommended index and offers several recommended parking fences. The lower the score, the more recommended parking. And the system will provide detailed information to users, including DBS flows of the recommended parking fences, density, active level, distances, and whether to cross the road. Different discounts can also further be developed based on different guidance schemes to encourage users to accept the guidance actively.

$$\begin{aligned} R_m = {} & w_o \cdot zscore\,(O_m) + w_l \cdot zscore\,(L_m) \\ & + w_a \cdot zscore\,(A_m)\,, \end{aligned} \tag{11}$$

where $zscore\,(\cdot)$ is the standardized equation; R_m is the recommended index for the M-th fence; w_o, w_l, w_a are the comprehensive index weight of the electric fence, the distance weight between the bicycle and the recommended fence, and the active level weight, respectively. In this work, they are 0.3, 0.5, 0.2.

After guiding the user to park, DBS dynamic guidance system updates the value of all fences flow and bicycle locations. Finally, we obtain the rebalanced fences for the current urban area. Algorithm 2 gives the details about DBS dynamic guidance method.

V. EXPERIMENTS

In this section, we conducted experiments on real-world datasets to verify the effectiveness and accuracy of DBSDR system. We collected DBS order datasets, DBS trajectory records, and electric fence datasets in Xiamen city from December 21 to 25, 2020. There are 12 million trajectory data (recorded once in 15 seconds), 600,000 order data, and 14071 fences.

A. DATASETS

The DBS trajectory dataset was collected from the road network in Xiamen city, China, containing the bicycle ID, source of bicycle, longitude, latitude, and its timestamp. Each electric fence data contains the fence ID and its position information. Bicycle-sharing order datasets are routinely collected from all stationary bicycles. Each order data contains the bicycle ID, longitude, latitude, the locking status of bicycles (i.e., opening and closing), and an updated timestamp of the locking status. Due to the inaccurate coordinate positioning of some bicycles, we preprocessed the data with data cleaning and filtering to obtain valid datasets.

B. CASE STUDY

To better understand the variation and trend of the early peak tidal phenomenon, the geohash is employed to encode the bicycles and the electric fences. Orders far from the fence

Algorithm 2 DBS Dynamic Guidance

Input:

X : The bicycle dataset and the electric fence dataset; Y: The bicycles in overload fences;

δ : The distance threshold for acceptable guiding distance;

Output:

The rebalanced fences.

for each bicycle in Y and each fence in X **do**
    geohash encode each bicycle and each fence and matches them with nine grids;
    extract the neighbor fences from the nine grids near each bicycle;
    remove or refuse bicycles parking without neighbor fences;
    for each fence in neighbor fences **do**
    calculate the shortest distance L_m by Dijkstra algorithm;
        if $L_m < \delta$ **then**
            obtain the nearest fence geohash encoding G (including multiple fences);
        end if
    end for
end for
for each fence in G **do**
    calculate the recommended index R_m of each fence using Equation (11);
    sort index and show L_m, whether cross the road, P_m, D_m, A_m by Equation (5) (7) (10);
    if the users accept the guidance **then**
        target fences flow $+1$, original fences flow -1;
        update the value of all fences flow;
    end if
end for
return rebalanced fences.

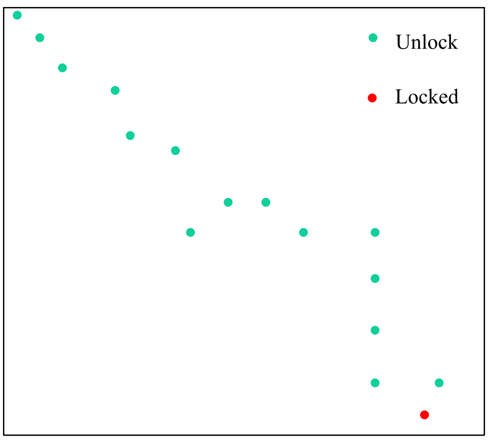


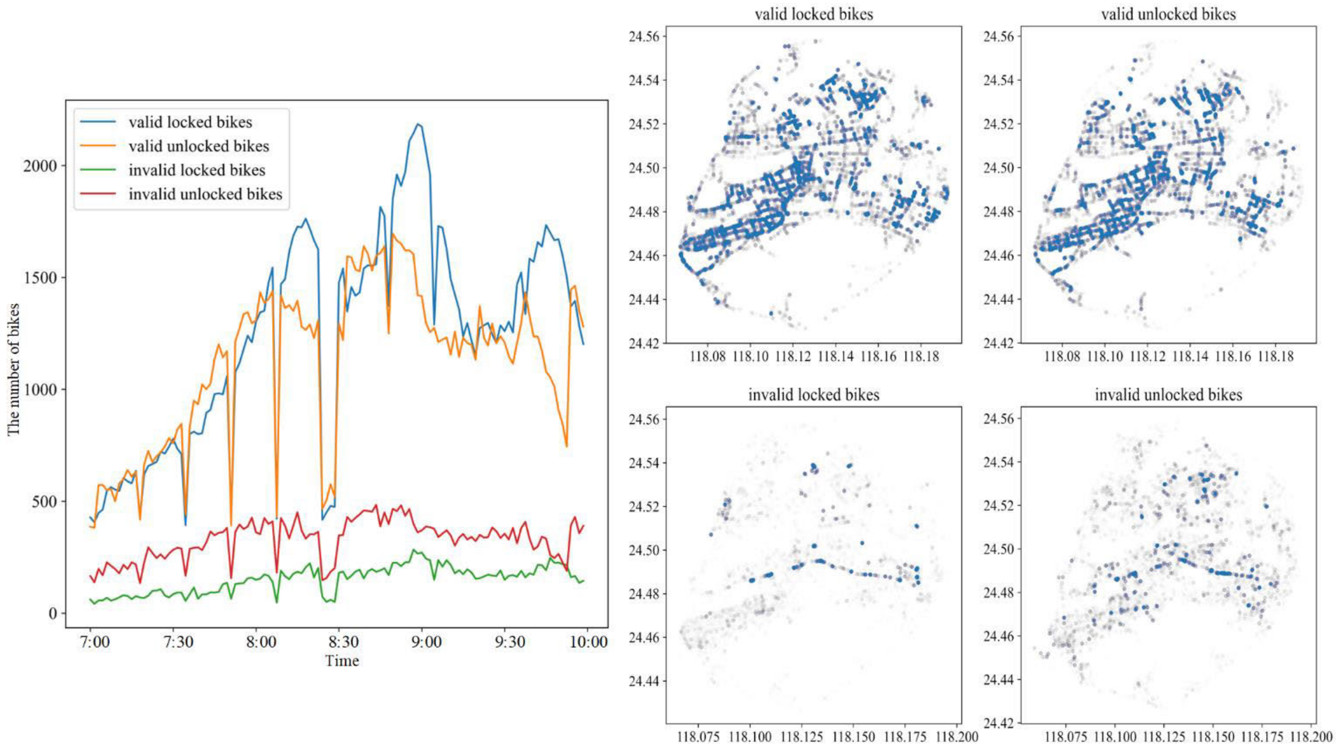
FIGURE 3. The error states bicycle-sharing.

were selected as outliers and ongoing situations. In general, order states have two states, incuding opening and closing. Due to equipment failure, the same bicycles may appear opening many times and the closing status are not equal to the locking state. In Figure 3, the bicycles move continuously when the unlocking data occur. Therefore, we need to retain the data generated when the bicycle state changes. For continuous unlocking data, only the first is retained; for continuous lock data, only the last one is retained.

In this work, too long geohash codes directly match the electric fences can cause some electric fences to be not completely covered. Therefore, we use 7-bit geohash code that is 153 * 153 square lattice for four vertex positioning that can cover 306 * 306 fence. Bicycles that cannot be covered only account for less than 1 % of the total amount of data, which can be classified as "disorderly parking". The spatio-temporal distribution of valid orders and invalid orders is shown in Figure 4. There is an obvious tidal phenomenon in the time variation trend between valid orders and invalid orders. In the space of valid orders and invalid orders, the "disorderly parking" phenomenon in some areas is more serious.

Figure 5 shows the tidal zones distribution under three different indicators. Figure 5(a) shows the tidal distribution using the number of bicycle-sharing P_m as an indicator. It shows the top 40 crowded tidal zones which contain 559 electric fences, showing the phenomenon of "large number of bicycles". Figure 5(b) shows the tidal distribution using the parking demand density D_m as an indicator. It shows the top 40 crowded tidal zones which contain only 64 electric fences, showing the phenomenon of "bicycle congestion". Most of the fences with high density were observed to be clustered separately. Figure 5(c) shows the tidal distribution according to the comprehensive index O_m. It shows the top 40 crowded tidal zones which contain 353 electric fences, showing the phenomenon of "much and congestion".

The tidal zones analysis and scheduling feasibility under the comprehensive index O_m of Figure 5(c) are discussed.



(a) Temporal distribution of bicycle

(b) Spatial distribution of bicycle

FIGURE 4. The spatio-temporal distribution of valid and invalid orders. There is an obvious tidal phenomenon in the time variation trend between valid orders and invalid orders. In the space of valid orders and invalid orders, the "disorderly parking" phenomenon in some areas is more serious.

(a) The phenomenon of 'large number of bicycles'

(b) The phenomenon of 'bicycles congestion'

(c) The phenomenon of 'much and congestion'

FIGURE 5. The tidal zone distribution under three different indicators. (a) The tidal distribution using the number of bicycle-sharing P_m as an indicator. (b) The tidal distribution using the parking demand density D_m as an indicator. (c) The tidal distribution according to the comprehensive index O_m.

Figure 5(c) shows that these tidal zones are central business districts (CBDs), schools, and commercial areas, which are in line with the actual situation. The overload fences are obtained by the comprehensive indicator O_m. The minimum congestion degree in the overload fences is (five-day retention flow 27, retention density 1.51, comprehensive score = 1.616992). Figure 6 shows details of the 353 fences in the top 40 tidal zones, and it could be found that 52.4 % of the 353 fences are not overload fences while the remaining 47.6 % fences contains 92.9 % (14,766) of bicycles.

A tidal zone is selected for observation to prove whether there is a large tidal degree gap between adjacent electric fences. Figure 7 shows a tidal zone containing 35 electric fences. The overload fences are marked as red and the idle fence is marked as blue. It is found that there is a great degree gap in the adjacent electric fences, which fully shows that

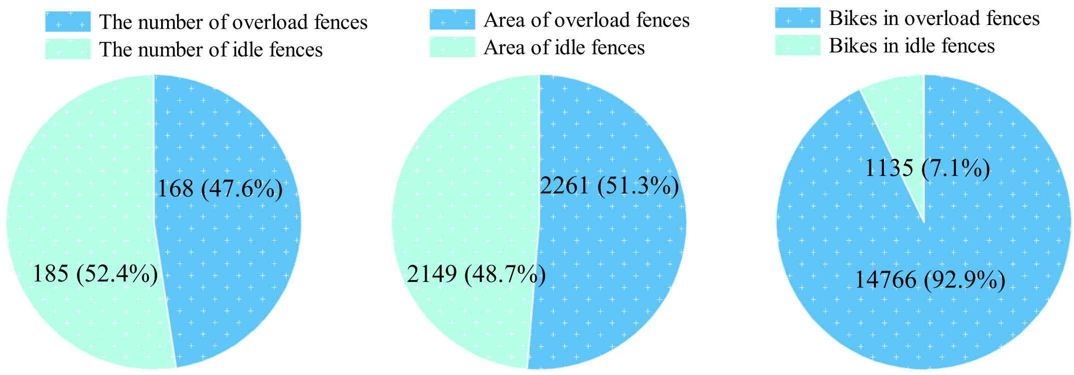


FIGURE 6. The division of 353 fences in the top 40 tidal zones.

FIGURE 7. A tidal zone containing 35 electric fences.

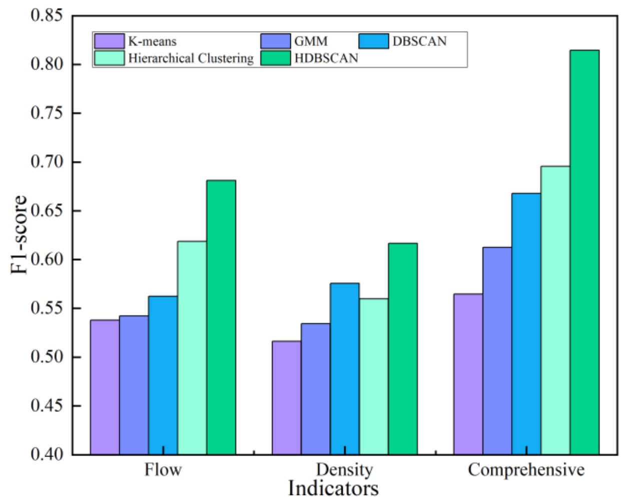


FIGURE 8. Performance Comparison of Different Clustering Methods.

there are some relatively idle fences in the tidal zones for guidance, and provides feasible support for the guidance on parking management schemes.

C. COMPARISON WITH OTHER CLUSTERING ALGORITHMS

In this section, the HDBSCAN-based tidal zones clustering algorithm will be discussed in the comparison with the traditional DBSCAN, K-means, Hierarchical Clustering and GMM algorithm. We use the following metrics to evaluate the clustering result accuracy:

$$precision = \frac{|TP|}{|TP| + |FP|}, \tag{12}$$

$$recall = \frac{|TP|}{|TP| + |FN|}, \tag{13}$$

$$F1_{score} = \frac{2 \cdot precision \cdot recall}{precision + recall}. \tag{14}$$

As shown in Figure 8 and Table 3, we analyze the accuracy of different clustering methods under three indicators. The clustering results show that the HDBSCAN-based tidal zones clustering algorithm outperforms the baseline methods in three indicators. Specifically, the DBS density of the fences with the same flow will be very different and single density cannot reflect the flow of the fences, which lends to the clustering results on a composite indicator O_m are better than single-flow indicators P_m and density indicator D_m. Furthermore, when a disconnected nonconvex set appears, K-means and GMM will have an incorrect classification. When the same cluster has a little of sparse points, DBSCAN divides points in the same cluster into multiple clusters. Hierarchical Clustering combines multiple cluster crossing points into the same cluster.

In contrast, the HDBSCAN-based tidal zones clustering algorithm combines DBSCAN and Hierarchical Clustering, which can be used to find clusters of multiple shapes, and will not split the cluster wrongly when there are a little of discontinuous points. The clustering results show that the optimal performance of the HDBSCAN-based tidal zones clustering algorithm can reach 81.45%.

To further explore the effectiveness of the proposed DBSDR system, we analyzed the guiding strategies of overload fences in four tidal zones including guiding distance and DBS flow of target fences. The two target fence

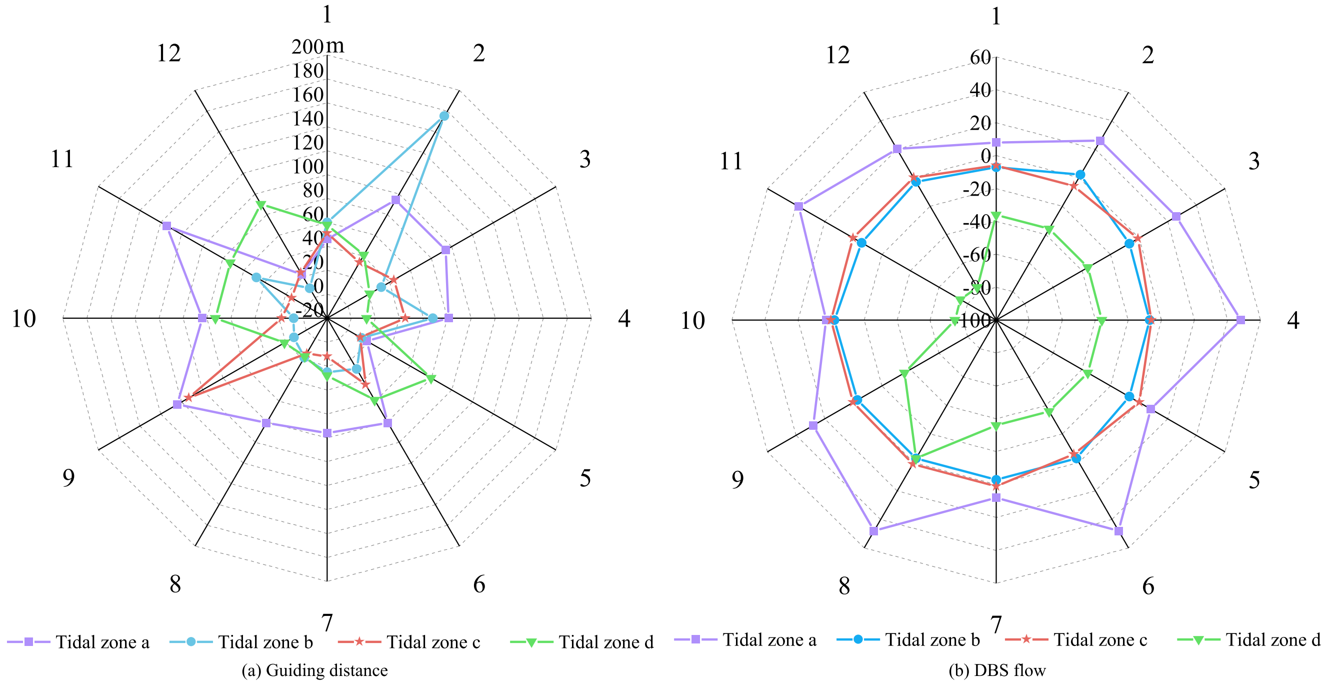


FIGURE 9. The two target fence indicators of different overload fences in four tidal zones. (a) The guiding distance of different overload fences in four tidal zones, most of which is within 100 meters. For users, it could be generally accepted. (b) The DBS flow of target fences of different overload fences in four tidal zones, most of the target fences are in the absence of DBS, which just solves the imbalance problem between overload fences and idle fences. The DBS flow of the target fences in the tidal zone a is slightly higher but also belongs to idle fences.

TABLE 3. Descriptions of comparing clustering algorithm.

No.	Algorithm	Description
1	K-means	The distance-based clustering algorithm combines classics and simplicity.
2	Gaussian Mixture Model (GMM)	The clustering method by selecting components to maximize posterior probability.
3	DBSCAN	The density-based clustering algorithm can find clusters of arbitrary shape in the noisy spatial database.
4	Hierarchical Clustering	The clustering method by creating a hierarchical nested tree can calculate the similarity of different points.
5	HDBSCAN	The new clustering method combines the analytic hierarchy process and density clustering.

indicators of different overload fences are shown in Figure 9. Figure 9(a) shows the guiding distance of different overload fences in four tidal zones, most of which is within 100 meters. For users, it could be generally accepted. On the other hand, Figure 9(b) shows that the DBS flow of target fences of different overload fences in four tidal zones, most of the target fences are in the absence of DBS, which just solves the imbalance problem between overload fences and idle fences. The DBS flow of the target fences in the tidal zone a is slightly higher but also belongs to idle fences.

D. PARAMETER SETTINGS

The experimental server was equipped with an Intel i5 CPU, NVIDIA 3060 GPU. The version of Python is 3.7. The parameter min_cluster_size is set to 3, gen_min_span_tree is True, and cluster_selection_epsilon is 0.0003.

VI. CONCLUSION

In this work, we propose a Dockless Bicycle-sharing Dynamic Rebalance (DBSDR) system to dynamically provide optimal bicycle guidance for the DBS network. The DBSDR system consists of DBS tidal zone identification, evaluation framework of overload fences, and DBS dynamic guidance system. In detail, tidal zone identification and location from each fence with bicycle flows are provided based on HDBSCAN clustering. Then, both DBS flow and parking demand density are considered in the model's evaluation framework. Experimental results on real-world datasets validated the effectiveness and accuracy of DBSDR system. For the DBS network, DBS dynamic guidance system provides users with five target fence indicators that can effectively guide users to balance DBS for solving tidal phenomenon, and the method could also be applied to the scheduling of DBS in various regions. For the society, traffic rules do not allow disorderly parking situation and punish users who park disorderly, the system can assist the normal operation of traffic. Meanwhile, the system will provide detailed information to users, including DBS flows of the recommended parking fences, density, active level, distances, and whether to cross the road. Users can choose their own parking fences. In the future, different discounts can also further be developed based on different guidance schemes to encourage users to accept the guidance actively. We will further study the discounts that encourage user scheduling. In addition, bicycle-sharing station dynamic planning is also worthy of further study.

REFERENCES

[1] M. Hua, X. Chen, S. Zheng, L. Cheng, and J. Chen, "Estimating the parking demand of free-floating bike sharing: A journey-data-based study of Nanjing, China," *J. Cleaner Prod.*, vol. 244, Jan. 2020, Art. no. 118764.

[2] X. Yuan, J. Senlin, Z. Chongxia, L. Wei, H. Yan, and Y. Hejun, "High accuracy virtual electronic fence management technique based on GNSS," in *Proc. 13th IEEE Int. Conf. Electron. Meas. Instrum. (ICEMI)*, Oct. 2017, pp. 79–83.

[3] Q. Yu, Y. Xie, W. Li, H. Zhang, X. Liu, W.-L. Shang, J. Chen, D. Yang, and J. Yan, "GPS data in urban bicycle-sharing: Dynamic electric fence planning with assessment of resource-saving and potential energy consumption increasement," *Appl. Energy*, vol. 322, Sep. 2022, Art. no. 119533.

[4] Y. Zhang, D. Lin, and Z. Mi, "Electric fence planning for dockless bike-sharing services," *J. Cleaner Prod.*, vol. 206, pp. 383–393, Jan. 2019.

[5] L. Caggiani and R. Camporeale, "Toward sustainability: Bike-sharing systems design, simulation and management," *Sustainability*, vol. 13, no. 14, p. 7519, Jul. 2021.

[6] Q. Jiang, L.-C. Chen, and W. Wei, "Service design consideration of solving the problem of disorderly-parked unbundled sharing bicycles with low-carbon policy orientation," in *Proc. IEEE Int. Conf. Appl. Syst. Invention (ICASI)*, Apr. 2018, pp. 524–527.

[7] Y. Hui, Y. Hui, Y. Xie, Q. Yu, X. Liu, and X. Wang, "Hotspots identification and classification of dockless bicycle sharing service under electric fence circumstances," *J. Adv. Transp.*, vol. 2022, Mar. 2022, Art. no. 5218254.

[8] Z. Xiao, H. B. Lim, and L. Ponnambalam, "Participatory sensing for smart cities: A case study on transport trip quality measurement," *IEEE Trans. Ind. Informat.*, vol. 13, no. 2, pp. 759–770, Apr. 2017.

[9] Y. Zhang and Z. Mi, "Environmental benefits of bike sharing: A big data-based analysis," *Appl. Energy*, vol. 220, pp. 296–301, Jun. 2018.

[10] N. Gast, G. Massonnet, D. Reijsbergen, and M. Tribastone, "Probabilistic forecasts of bike-sharing systems for journey planning," in *Proc. 24th ACM Int. Conf. Inf. Knowl. Manage.*, Melbourne, VIC, Australia, Oct. 2015, pp. 703–712.

[11] Y. Li, Y. Zheng, H. Zhang, and L. Chen, "Traffic prediction in a bike-sharing system," in *Proc. 23rd SIGSPATIAL Int. Conf. Adv. Geograph. Inf. Syst.*, Seattle, WA, USA, Nov. 2015, p. 33.

[12] Y. Duan and J. Wu, "Optimizing rebalance scheme for dock-less bike sharing systems with adaptive user incentive," in *Proc. 20th IEEE Int. Conf. Mobile Data Manage. (MDM)*, Jun. 2019, pp. 176–181.

[13] A. Audikana, E. Ravalet, V. Baranger, and V. Kaufmann, "Implementing bikesharing systems in small cities: Evidence from the Swiss experience," *Transp. Policy*, vol. 55, pp. 18–28, Apr. 2017.

[14] T. Fukushige, D. T. Fitch, and S. Handy, "Can an incentive-based approach to rebalancing a dock-less bike-share system work? Evidence from Sacramento, California," *Transp. Res. A, Policy Pract.*, vol. 163, pp. 181–194, Sep. 2022.

[15] Y. Ma, X. Qin, J. Xu, and W. Wang, "A hierarchical public bicycle dispatching policy for dynamic demand," in *Proc. IEEE Int. Conf. Service Oper. Logistics, Informat. (SOLI)*, Jul. 2016, pp. 162–167.

[16] L. Jia, D. Yang, Y. Ren, Q. Feng, B. Sun, C. Qian, Z. Li, and C. Zeng, "An electric fence-based intelligent scheduling method for rebalancing dockless bike sharing systems," *Appl. Sci.*, vol. 12, no. 10, p. 5031, May 2022.

[17] J. Song, L. Zhang, Z. Qin, and M. A. Ramli, "A spatiotemporal dynamic analyses approach for dockless bike-share system," *Comput., Environ. Urban Syst.*, vol. 85, Jan. 2021, Art. no. 101566.

[18] H. Xu, F. Duan, and P. Pu, "Dynamic bicycle scheduling problem based on short-term demand prediction," *Appl. Intell.*, vol. 49, no. 5, pp. 1968–1981, May 2019.

[19] Q. Chen, M. Liu, and X. Liu, "Bike fleet allocation models for repositioning in bike-sharing systems," *IEEE Trans. Transp. Syst. Mag.*, vol. 10, no. 1, pp. 19–29, Spring 2018.

[20] C. Chen, K. Li, S. G. Teo, X. Zou, K. Wang, J. Wang, and Z. Zeng, "Gated residual recurrent graph neural networks for traffic prediction," in *Proc. AAAI Conf. Artif. Intell.*, 2019, pp. 485–492.

[21] G. Peters, F. Crespo, P. Lingras, and R. Weber, "Soft clustering—Fuzzy and rough approaches and their extensions and derivatives," *Int. J. Approx. Reasoning*, vol. 54, no. 2, pp. 307–322, Feb. 2013.

[22] L. Chen, D. Zhang, L. Wang, D. Yang, X. Ma, S. Li, Z. Wu, G. Pan, T.-M.-T. Nguyen, and J. Jakubowicz, "Dynamic cluster-based over-demand prediction in bike sharing systems," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, Heidelberg, Germany, Sep. 2016, pp. 841–852.

[23] J. Huang, X. Wang, and H. Sun, "Central station based demand prediction in a bike sharing system," in *Proc. 20th IEEE Int. Conf. Mobile Data Manage. (MDM)*, Jun. 2019, pp. 346–348.

[24] E. A. A. Alaoui and S. C. K. Tekouabou, "Intelligent management of bike sharing in smart cities using machine learning and Internet of Things," *Sustain. Cities Soc.*, vol. 67, Apr. 2021, Art. no. 102702.

[25] P. Xie, T. Li, J. Liu, S. Du, X. Yang, and J. Zhang, "Urban flow prediction from spatiotemporal data using machine learning: A survey," *Inf. Fusion*, vol. 59, pp. 1–12, Jul. 2020.

[26] R. Ying, R. He, K. Chen, P. Eksombatchai, W. L. Hamilton, and J. Leskovec, "Graph convolutional neural networks for web-scale recommender systems," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, London, U.K., Jul. 2018, pp. 974–983.

[27] Y. Wu, H. Cao, G. Yang, T. Lu, and S. Wan, "Digital twin of intelligent small surface defect detection with cyber-manufacturing systems," *ACM Trans. Internet Technol.*, early access, Nov. 17, 2022, doi: 10.1145/3571734.

[28] J. Gou, L. Sun, B. Yu, S. Wan, W. Ou, and Z. Yi, "Multi-level attention-based sample correlations for knowledge distillation," *IEEE Trans. Ind. Informat.*, early access, Sep. 26, 2022, doi: 10.1109/TII.2022.3209672.

[29] Y. Zhang, F. Zhang, Y. Jin, Y. Cen, V. Voronin, and S. Wan, "Local correlation ensemble with GCN based on attention features for cross-domain person Re-ID," *ACM Trans. Multimedia Comput., Commun., Appl.*, early access, Jun. 9, 2022, doi: 10.1145/3542820.

[30] Y. Wu, H. Guo, C. Chakraborty, M. Khosravi, S. Berretti, and S. Wan, "Edge computing driven low-light image dynamic enhancement for object detection," *IEEE Trans. Netw. Sci. Eng.*, early access, Feb. 14, 2022, doi: 10.1109/TNSE.2022.3151502.

[31] J. Chen, K. Li, K. Li, P. S. Yu, and Z. Zeng, "Dynamic planning of bicycle stations in dockless public bicycle-sharing system using gated graph neural network," *ACM Trans. Intell. Syst. Technol.*, vol. 12, no. 2, pp. 1–22, Apr. 2021.

[32] M. Liu and X. Xu, "Dockless bike-sharing reallocation based on data analysis: Solving complex problem with simple method," in *Proc. IEEE 3rd Int. Conf. Data Sci. Cyberspace (DSC)*, Jun. 2018, pp. 445–450.

[33] P. Leclaire and F. Couffin, "Method for static rebalancing of a bike sharing system," *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 1561–1566, 2018.

[34] J. Liu, L. Sun, W. Chen, and H. Xiong, "Rebalancing bike sharing systems: A multi-source data smart optimization," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, San Francisco, CA, USA, Aug. 2016, pp. 1005–1014.

[35] J. Schuijbroek, R. C. Hampshire, and W.-J. van Hoeve, "Inventory rebalancing and vehicle routing in bike sharing systems," *Eur. J. Oper. Res.*, vol. 257, no. 3, pp. 992–1004, Mar. 2016.

[36] L. Shi, Y. Zhang, W. Rui, and X. Yang, "Study on the bike-sharing inventory rebalancing and vehicle routing for bike-sharing system," *Transp. Res. Proc.*, vol. 39, pp. 624–633, Jan. 2019.

[37] X. Zhang, H. Yang, R. Zheng, Z. Jin, and B. Zhou, "A dynamic shared bikes rebalancing method based on demand prediction," in *Proc. IEEE Intell. Transp. Syst. Conf. (ITSC)*, Oct. 2019, pp. 238–244.

[38] D. Zhang, C. Yu, J. Desai, H. Y. K. Lau, and S. Srivathsan, "A time-space network flow approach to dynamic repositioning in bicycle sharing systems," *Transp. Res. B, Methodol.*, vol. 103, pp. 188–207, Sep. 2017.

[39] Z. Tian, J. Zhou, W. Y. Szeto, L. Tian, and W. Zhang, "The rebalancing of bike-sharing system under flow-type task window," *Transp. Res. C, Emerg. Technol.*, vol. 112, pp. 1–27, Mar. 2020.

[40] Y. Ji, X. Jin, X. Ma, and S. Zhang, "How does dockless bike-sharing system behave by incentivizing users to participate in rebalancing?" *IEEE Access*, vol. 8, pp. 58889–58897, 2020.
[41] R. J. G. B. Campello, D. Moulavi, and J. Sander, "Density-based clustering based on hierarchical density estimates," in *Proc. Pacific-Asia Conf. Knowl. Discovery Data Mining*, 2013, pp. 160–172.

LYUCHAO LIAO (Senior Member, IEEE) received the Ph.D. degree in traffic information engineering and its control from Central South University, in 2015. He was a Postdoctoral Researcher at Tsinghua University, from 2016 to 2018, and visited the University of Essex, U.K., in 2019. Currently, he is a Full Professor with the School of Transportation, Fujian University of Technology (FJUT). His research interest includes big data and artificial intelligence in transportation, such as driving behavior analysis, traffic state prediction, and traffic road network optimization.

BEN LI is currently pursuing the master's degree with the School of Transportation, Fujian University of Technology (FJUT). His research interests include artificial intelligence, big data in transportation, traffic state prediction, intelligent transportation systems, and multimodal data analysis.

DEJUAN HUANG received the master's degree in transportation engineering from Central South University, in 2013. She is currently an Engineer with the School of Transportation, Fujian University of Technology (FJUT). Her research interests include data analysis and artificial intelligence in transportation.

ZHU XIAO (Senior Member, IEEE) received the M.S. and Ph.D. degrees in communication and information systems from Xidian University, China, in 2007 and 2009, respectively. From 2010 to 2012, he was a Research Fellow with the Department of Computer Science and Technology, University of Bedfordshire, U.K. He is currently a Full Professor with the School of Computer Science and Electronic Engineering, Hunan University, China. His research interests include wireless localization, the Internet of vehicles, and intelligent transportation systems. He is an Associate Editor for IEEE Transactions on Intelligent Transportation Systems.

QI ZHENG is currently pursuing the master's degree in information enginnering from the Fujian University of Technology (FJUT). Her research interests include artificial intelligence in transportation, reinforcement learning, and deep learning. Currently, she is delicated to traffic regulation.