




Article

A Hybrid Dispatch Strategy Based on the Demand Prediction of Shared Bicycles

Shu Shen ^{1,2,3,*}, Zhaoqing Wei ^{1,†}, Lijuan Sun ^{1,2,*}, Khalida Shaheen Rao ¹, and Ruchuan Wang ^{1,2}

¹ School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210003, China; 1017041027@njupt.edu.cn (Z.W.); K.shaheenrao@gmail.com (K.S.R.); wangrc@njupt.edu.cn (R.W.)

² Jiangsu High Technology Research Key Laboratory for Wireless Sensor Networks, Nanjing 210003, China

³ Tongda College, Nanjing University of Posts & Telecommunications, Yangzhou 225127, China;

* Correspondence: shens@njupt.edu.cn (S.S.); sunlijuan_nupt@163.com (L.S.); Tel.: +86-25-8349-2470 (S.S. & L.-J.S.)

† These authors contributed equally to this work.

Received: 4 March 2020; Accepted: 12 April 2020; Published: 16 April 2020



Abstract: With the advent of pile-less shared bicycles, the techniques initially used for public bicycle dispatching were unable to fulfill the routine dispatch tasks, resulting in constant bicycle crowding. In this paper, to alleviate the mess of shared bicycles, we propose a hybrid dispatching algorithm based on bicycle demand data. We take the bicycle stations' imbalance as an optimization index and use greedy ideas to ensure that after each dispatch all stations get the smallest imbalance. In addition, it is suggested that two assessment metrics evaluate the efficiency of the dispatching technique from the users and operators' perspectives. It is shown that the proposed dispatching algorithm performs better in terms of user satisfaction and operator revenue, and is less affected by bicycle distribution compared with the traditional manual scheduling algorithm.

Keywords: IoSB; demand forecast; ARMA; bicycle dispatching; greedy strategy

1. Introduction

With the development and application of advanced technologies such as communication and microelectronics, the concept of smart city arises at the right moment. Smart city has a much broader definition, which includes almost every aspect in human beings' life, e.g., smart medical, smart energy, smart traffic, and smart home. As a form of intelligent transport, the shared bicycle, also known in some references as bicycle-sharing [1–4], is about 55 years old and has four generations. The shared bicycle is thriving in smart cities especially in recent years and most of them have been successful. There are three clear benefits on the shared bicycle [5], as follows:

- It is an utter green path that can reduce environmental effects, regulate air pollution, and effectively save money. There are no exhaust emissions that are dangerous for the atmosphere without fossil fuel consumption. Since cycling is a form of physical practice, both the community and users benefit from the shared bicycle.
- For the last mile journey, it provides the perfect solution. Actually, on long-distance journeys, many citizens do not prefer cycling today. However, given that, in most instances, the range of metro or bus stations and commercial or residential communities is not that long, shared bicycles are most cost-efficient and suitable for people for short distances.
- The urban market is a typical example of public resource-shared operation, capable of improving the cycling usage rate. Today, increasing numbers of people favor shared bicycles than their conventional private bicycles. The traffic pressure will, therefore, be helpful.

Referring to [5], Shen et al. first divided the development process of the shared bicycle into four phases, and suggested the Internet of Shared Bicycle (IoSB) concept. With the advent of many new technologies such as QR Code or LTE, the fourth-generation connected bicycle [6,7], called one form of intelligent hardware, has evolved so rapidly over the past few years. Most significantly, the change from the pile-parking mode to the pile-less parking mode of shared bicycles is a new attempt and also a double-edged sword. Great convenience comes at the expense of unnecessary supply of shared bicycles which usually manifests as too many bicycles being left unused or too few bicycles to meet demands. Insufficient timely and reasonable bicycle scheduling trigger two extreme situations above and may affect the social order.

To address and resolve the above-mentioned problems, this paper proposed a hybrid dispatch scheme based on demand forecasting data. This hybrid dispatch system combines manual dispatch and user reward dispatch. Firstly, according to the travel records of Mobike, we divided the city into many travel zones. Afterward, we pick one of the travel areas for study in this paper and choose a certain number of bicycle travel points in the research area using the clustering technique. Instead, at the next time of the mutual bicycle travel point, we use Autoregressive Moving Average (ARMA) to predict bicycle demand. Subsequently, referring to the travel point forecast data, we established the manual and user incentive hybrid dispatch strategy based on the greedy strategy to ensure the lowest degree of equilibrium of all bicycle travel points in the study area after each dispatch. Finally, we take user satisfaction and total revenues of shared bicycle operators as evaluation indicators and carry out a simulation analysis of the hybrid dispatching strategy proposed in this paper.

The key contents of this paper are structured as follows: Section 2 presents some study directions concerning the shared bicycle at home and abroad; Section 3 illustrates the entire shared bicycle dispatching scenario and framework; Section 4 introduces the concept of manual dispatch and user incentive dispatching suggested in this paper respectively. Section 5 describes the setting of experimental parameters, input data, and experimental methods, and analyzes the results of the experiment. Lastly, Section 6 provides conclusions and future work.

2. Related Work

Since the advent of the shared bicycle, many scholars around the world have been studying relevant questions. The authors in [5] suggested the IoSB concept, which includes the perception, communication, application and safety layers, and showed five-layer model IoSB architecture. In IoSB, the transition from the original pile-shared bicycle to the current pile-less sharing bicycle has many problems to overcome, such as environmental information collection, node location, low power communication technology, bicycle dispatching scheme, path management, etc.

IoSB is essentially a type of Wireless Sensor Networks (WSNs), and its information collection has long been related to research. In [8], Sha et al. proposed a type of low-latency data gathering method with multi-sink in WSNs. Network division, multi-sink, efficient sleep scheduling and the sensing radius adjustment strategies are utilized to reduce the redundancy in network coverage and further improve the energy efficient.

With its advantages of low energy consumption, node positioning technology in the WSN network also has certain reference significance for IoSB. In [9], a type of energy-efficient target tracking approach based on grids in WSNs was proposed. The concept of virtual grids and its three kinds of states including tracking state, prepared-tracking state, and preparing-tracking state were defined to reduce energy consumption and enhance the accuracy of node localization.

Reliable data transmission is an important guarantee for data-based applications, especially real-time ones such as a shared bicycle. Communication efficiency is determined by the technical norm, network topology, base station antenna sensitivity, etc. Here, are some of the communications technologies that are commonly used in mobile networks, like the shared bicycle. LTE is the technical standard that all existing smartphones adopt, enabling data terminals to have a faster speed of network transmission which is ubiquitous and suitable for wide-range, high-bandwidth applications like

IoSB [10]. However, its modules' cost and energy consumption are too high to be applied in IoSB. The Low-Power Wide-Area Network (LPWA) has strong penetration and is designed for applications in complex urban environments with long communication distances [11]. Compared to existing LTE networks, NB-IoT, as a kind of LPWA technique, has the advantages of wide coverage, low power consumption, low module costs, and a large connection [12], and has been used in applications such as water quality monitoring [13], smart parking [14], smart hospital [15], and shared bicycles.

Shared bicycle dispatching is crucial to rebalancing bicycles and can be separated into static dispatch and dynamic dispatching. Static dispatching is to research the cumulative mutual bicycle dispatch distance as a measure of optimization. In comparison, dynamic dispatching is based on operator revenue or user riding satisfaction optimization. The dispatching route problem between static leasing sites, which belongs to static dispatching, is studied in [16]. In conclusion, the problem of shared bicycles' static dispatching can be considered as a problem of path optimization, which can generally be modeled using linear integer programming [17–19] and found using heuristic algorithms such as Genetic Algorithm (GA) [20], Particle Swarm Optimization (PSO) [21], Ant Colony Optimization (ACO) [22], Simulated Annealing Algorithm (SAA) [23], and so on. Dynamic dispatching pays more attention to the current state or possible future state of the dispatching area and the surrounding region compared with static dispatching and then makes better, more effective and far-sighted scheduling decisions based on the current time and future state data on the map. There are a lot of cases on the demand forecasting of the shared bicycle. For example, in [24], the author proposed a moment-based hybrid prediction method. In that work, a fuzzy mean genetic algorithm and backpropagation network was used to predict the future demand of a certain region of the shared bicycle. Finally, the proposed model was verified by using the actual trip data of well-known foreign public bicycle operators. Bicycle demand in a certain area is related to many factors, such as the distribution of infrastructure in the region to be forecasted [25,26], ambient temperature, and sunny/rainy conditions [27], unexpected large events like holidays or strikes [28,29], time period [30], etc.

On the other hand, the mining of data on the bicycle track supports urban design and planning. In [31], MSRA suggested an estimated greedy-based solution to suggest the bike lane planning system, based on Mobike's monitor big data. In [32], the All-Pair Network Trajectory Similarity (APNTS) is studied to identify major routes of the bicycle for public transport. Additionally, a shared bicycle can be used to complement and connect with other public transport tools. In [33], the author used the mutual bicycle journey track data to refine the city's metro operating strategy.

3. Models and Preliminaries

3.1. System Models

Unlike other public transport carriers, there are no fixed stations for shared bicycles and their location is influenced by the tidal movement of users. Because of the complexity and mass of cycling, the distribution of shared bicycles is unbalanced and is difficult to rebalance, particularly for the shared bicycle with no piles. The importance of a high-efficiency strategy for bicycle dispatching is as follows:

- Reduce scheduling work and time;
- Alleviate cycling mess;
- Boost users' satisfaction with using shared bicycles.

Consider a scenario of shared bicycle system, as shown in Figure 1. With no fixed pile for parking, we conducted a cluster analysis on all cycling travel points, and extracted the central point of the cluster as the travel station that was illustrated by dots in Figure 1.

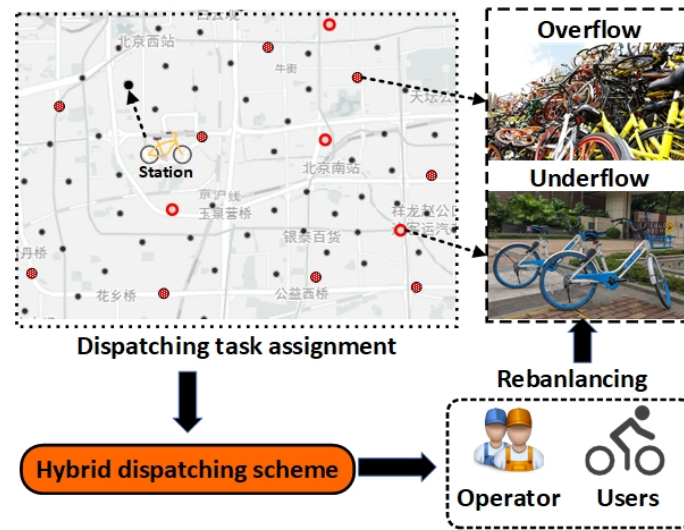


Figure 1. The dispatching model.

The travel stations in Figure 1 are modeled as the vertexes in an undirected graph $G = (V, E)$, where V and E represent all vertex sets and edge sets in undirected Graph G , respectively ($|E| \geq 1$). The E in graph G represents the paths between the shared bicycle stations, and the weight of E represents the relationship between the two shared bicycle points connected with them.

In addition, a two-tuple $S_i(t) = (num_i(t), pred_i(t))$ is used to describe the status of station i at time $t (i \in [0, |V| - 1])$. The station status S includes two parts:

- The number of bicycles initially in station i at time t is defined as $num_i(t)$;
- The Demand of station i in next period predicted at time t is defined as $pred_i(t)$.

To further describe the imbalance of station i at time t , we introduce a new parameter $imb_i(t) = pred_i(t) - num_i(t)$. The unbalanced station in G are divided into two types: overflow station ($imb_i(t) < 0$) and underflow station ($imb_i(t) > 0$). Next, in order to alleviate unbalanced stations, we assign the dispatching task to operator and users via hybrid dispatching scheme proposed in Section 4 to allocate the excess bicycles in overflow station to other stations that are underflow.

Firstly, we delegated some overcrowded dispatching tasks that needed to be dispatched to the operator immediately, which are called manual dispatching. Manual dispatching is direct and fast, which belongs to a macro level of dispatching scheme. Here, we introduce a parameter Max to describe the overflow threshold of the station, which means when $num_i(t) \geq Max$, the station i needs to be dispatched immediately by the operator. More details about manual dispatching are shown in Section 4.

Secondly, we give users a small number of less urgent dispatching tasks and pay them back rewards recognizing user flexibility and confusion, which are called user incentive dispatching and is a microscopic dispatch. Whether the user recognizes the tasks of dispatching is linked to the circumstance, time, environment, and other conditions of the user. After accepting the dispatch task, the user can choose a midpoint near his destination to park the bicycle, whose midpoint is here called *detour_node*. Taking into account the distance of the user walks, we select *detour_node* from the adjacent node set of the user's final destination. The adjacent node set of station i is defined as $adj^{(i)}$. Node set $adj_low^{(i)}(t)$ represents the node set in the adjacent nodes of station i where $imb_i(t) > 0$, and is defined as $adj_low^{(i)}(t) = \{n | num_n(t) \leq pred_n(t), n \in adj^{(i)}\}$.

Finally, we create a discrete-time simulation system to analyze the above hybrid dispatching strategy. In our simulation system, trip record data work as the input data and are formed as a quad $Items = \{t, depart, arrive, if_detour\}$, where t is the moment when the trip record occurred, depart and arrive respectively represent the beginning and end of the trip record ($depart \neq arrive$). *if_detour*

refers to whether user in the trip record is willing to detour to station a around the end station i for a constant reward $cost_u$, where $a \in adj_low^{(i)}(t)$.

3.2. Main Idea

We then work out our strategy for dispatching task allocation after model definition and develop the corresponding performance evaluation method. First, some assumptions are given as follows regarding the dispatching model:

- It ignores the time gap between arrival and departure of a bicycle at time t , and assumes that the departure time of a bicycle at a travel station is unified at the beginning of time t , and then reach the destination at the end of time t , during which the bicycle is independent of all the stations.
- It assumes that, during manual dispatching, the bicycle leaves the travel point at the beginning of time t , and then arrives at the destination at the next moment.
- It assumes that the revenue and cost of the operator only come from three parts: the charge for bicycle riding, the cost for manual dispatching, and the cost for user incentive dispatching.
- It assumes that the dispatching of shared bicycles is only carried out within the research region, and will not involve places outside the research region.

The main idea of this paper is to use a reasonable dispatching task allocation strategy to rebalance as many stations as possible. In order to achieve the above goal, we introduce the concept of overall station imbalance **IMB**, which indicates the gap of bicycle's supply and demand within all stations and the smaller value represents the more balanced bicycle distribution in IoSB. The specific definition is as Formula (2):

$$IMB(t) = \frac{1}{N} \sum_{i=1}^N |imb_i(t)| \quad (1)$$

For optimizing the **IMB**, the most important one is the prediction of the demand at the next moment. Trip record is a piece of data that records the starting point, ending point, time and other information of a trip. In our model, a quad Items is utilized to represent each trip record. The truth is that the trip record is a type of time series, which is an important branch of statistics. In this paper, we decided to use the ARMA model to predict the demand at the station for its prediction efficiency and the stability of time series. In the ARMA model, time series Y_t is a linear function of random error term of its current and early period and its early period value, which can be expressed as:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + u_t - \theta_1 u_{t-1} - \dots - \theta_q u_{t-q} \quad (2)$$

Finally, in order to evaluate the performance of the final dispatching model, two indicators are defined, which are total revenue for operators **B** and Users' satisfaction α . The indicator **B** is defined as Formula (3), which represents the operator's net income:

$$B = c_1 \cdot cost_t - c_2 \cdot cost_m - c_3 \cdot cost_u \quad (3)$$

where c_1 is total number of cycling orders within the observation period $[0, T - 1]$; c_2 is the total number of manual dispatching; c_3 is the total number of user incentive dispatching. $cost_t$ is the revenue of operator for single trip; $cost_m$ is the total cost of manual dispatching; $cost_u$ is the total cost of user incentive dispatching.

Users' satisfaction α is used to describe the extent which cycling needs are met in IoSB. In our work, we calculate the users' satisfaction α via the value of the percent of successful rental in all trip records. When there is no bicycle at the station where the lease order occurs, the trip order is deemed to be failed.

4. Hybrid Dispatching Scheme

4.1. Manual Dispatching Scheme

In this work, manual dispatching is a macro dispatching method and can be applied primarily to the situation where the number of bicycles in a travel station $num_i(t) \geq Max$. The number of bicycles in manual dispatching can be defined as $k = \left\lceil \frac{num_i(t) - Max}{k_0} \right\rceil$, where k_0 refers to the most bicycles each manual dispatching can handle. When $k > k_0$, $\left\lceil \frac{k}{k_0} \right\rceil$ times dispatching is needed. There are two difficulties in manual dispatching:

- The first is to select a better dispatching endpoint in the dispatching region;
- The second is to choose the optimal route between the start and end of dispatching.

We transform the manual dispatching scenario into the following mathematical problem as shown in Figure 2. The weight w_{ij} between station i and j which represented by edges in Figure 2 is defined as Formula (4):

$$w_{ij} = \frac{l_{ij}}{imb_i + imb_j} + \beta \quad (4)$$

where l_{ij} is the distance between station i and j ($i \neq j$); imb_i or imb_j is respectively the imbalance of station i or j , which has been defined in the previous Section 3. The parameter β is set to a constant value to ensure that the weight w_{ij} is always greater than 0. In Figure 2, when the manual dispatching starts to work, the system will firstly scan the global trip station and select the optimal dispatching destination according to the strategy proposed in this paper. Then, the path with the minimum cost is selected based on the weight w_{ij} between the start and end of dispatching. In the dispatching process, according to the disequilibrium and other parameters of the station, the allocation of bicycles to each station in the route is carried out.

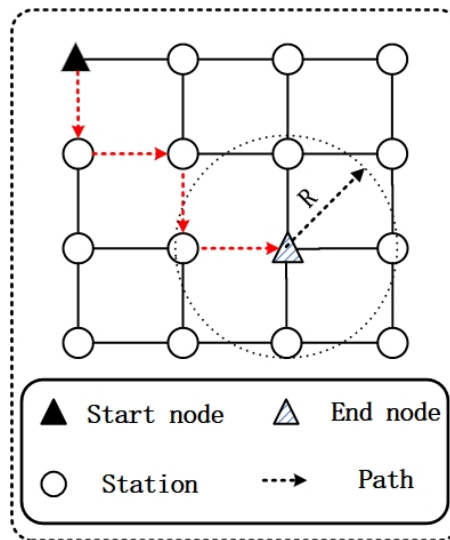


Figure 2. Manual dispatching scheme.

- (1) **The selection of dispatching endpoint:** As mentioned earlier in our work, we use greedy idea to ensure that the total imbalance of station at the next moment in the future within IoSB is minimized. Therefore, when selecting the endpoint in each manual dispatching, we can select the station with the most imbalance in IoSB, in order to achieve a rebalancing of the station. However, there is still a problem that needs to be considered in which bicycles in a single station will randomly move in the R field of the station, which makes it difficult to select the endpoint for manual dispatching. In the paper, we choose the sum of the imbalances of all stations in the range

of R at time t from a station as the basis for selecting the endpoint of manual dispatching, which is $imb_i^R(t)$ and its definition is shown in Formula (5):

$$imb_i^R(t) = \sum_R imb_i(t) \quad (5)$$

- (2) **The determination of dispatching path:** After solving the problem of dispatching endpoints, we need to figure out the route within manual dispatching. The truth is that the cost of manual dispatching is mainly related to the dispatching distance and the farther dispatching distance bring higher dispatching cost. Therefore, when seeking the optimal path, we give priority to the imbalance and distance of the stations. The weight w_{ij} in the undirected graph that is related to distance and imbalance of station is defined as the basis for route optimization. Finally, the Dijkstra algorithm is used to figure out the optimal path of manual dispatching. The details of manual dispatching are shown in Algorithm 1.

Algorithm 1 *Man_dispatch*(*start_station*, $S(t_n)$)

Input: The state information of all nodes in STG right moment $S(t)$, the station number to be dispatched *start_station*, the number of bicycles to be dispatched k .

Output: Dispatching path *Path*, bicycles assignment *Alloc*.

Initialize $Path = \{start_station\}$, $t = t_n$, $start_node = start_station$, $Alloc = \{\}$;

for each vertex i in G **do**

$imb_i^R(t_n) = \sum_R imb_i(t_n)$;

end for

$end_node = i'$, where $imb_{i'}^R(t_n) = \max_{i \in [0, N)} imb_i^R(t_n)$;

for each vertex i in G **do**

for each vertex j **do**

if $i \neq j$ **then**

$weight_{ij} = -\frac{imb_i}{l_{ij}}$;

end if

end for

end for

$Path = \text{Dijkstra}(start_node, end_node)$, $sum = \sum_{i \in Path} imb_i (imb_i > 0)$;

for each node i in *Path* **do**

if $imb_i \leq 0$ **then**

$Alloc \cup \{0\}$;

else

$Alloc \cup \{\lceil \frac{imb_i}{sum} * k \rceil\}$;

end if

end for

Return *Path*, *Alloc*.

4.2. User Incentive Scheme

The user incentive dispatching algorithm proposed in this paper is based on the idea of Crowdsourcing, which assigns the task of bicycle dispatching to various users of the shared bicycle, so as to achieve the goal of bicycle dispatching. The user incentive dispatching is a kind of micro-cycle dispatching with low cost and wide range, which is suitable for situation with a few bicycles and small dispatching scale. There are two difficulties in user incentive dispatching:

- The first is how to devise effective incentive strategy to encourage users to accept the Crowdsourcing dispatching task;
- The second is how to choose the target station of the detour to make the distribution of bicycle more balanced and reasonable.

For the first question, we do not consider it in the paper due to some complicated inducements like the personal condition, rental moment of bicycle, and the weather. The user incentive dispatching proposed in the paper is to make appropriate strategies to enable users to pay back their bicycles for a better choice and maximize the operator's revenue.

As shown in Section 4.2, we transformed user incentive dispatching into the following mathematical problems: A trip record shows the user going from *start_station* to *end_station*. In path1, the user travels directly from *start_station* to *end_station*. In path2, the user selects the detour station d_1 and then goes to *end_station*. In path3, the user selects the detour station d_2 and then goes to *end_station*. When the user chooses to detour, he will get a incentive $cost_u$ ($cost_u$ is assumed to be fixed here for the sake of system simplification). In Figure 3, we select all bicycle stations in the r field of the *end_station* in STG as the candidate detour station set $d = \{i \mid \|i - end_station\| \leq r\}$, where r is the maximum distance a user can walk when detouring. Then, we select the station which has maximum $pred_i(t)$ in the detour station set d as the detour node. The details of user incentive dispatching are shown in Algorithm 2.

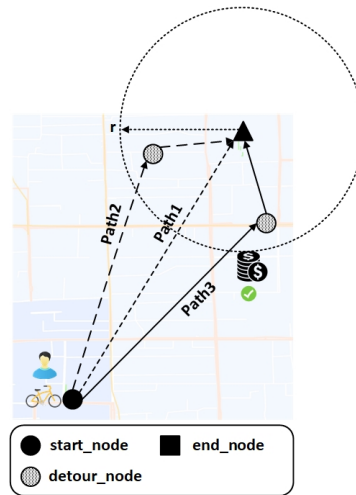


Figure 3. User incentive dispatching scheme.

Algorithm 2 *Man_dispatch*(*end_station*, $S(t_n)$)

Input: The state information of all nodes in STG right moment $S(t)$, the destination of the current user *end_station*.

Output: Recommended detour node for user.

Initialize $t = t_n$, *end_node* = *end_station*, *detour_node* = {*end_station*}, r

for each vertex i in G **do**

if $\|i - end_station\| \leq r$ && $i \neq end_station$ **then**
 detour_node $\cup \{i\}$;

end if

end for

candidate = *end_station*;

if $\max_{i \in detour_node} imb_i > 0$ **then**

candidate = $\max_{i \in detour_node} imb_i$;

end if

Return *candidate*.

5. Experiment

5.1. Experiment Preparation

In this section, we will evaluate our model in simulation to analyze the hybrid dispatching scheme. We first choose the trip data generated by Mobike's operation as source data which is given in an algorithm contest held by Mobike in 2017 [34]. In this paper, we only study and analyze the trip record data in the working day.

It is the procedure of discrete simulation system in Figure 4, where the function `Man_dispatch()` and `User_incented()` are respectively the manual dispatching algorithm and user incentive dispatching algorithm proposed in Section 4, and the function `Update_mapInfo()` is used to update the map status information after dispatching. Before the simulation, we need to filter the source data and preprocess its result. At the same time, we also need to initialize the station state S . When the simulation begins, the system first uses the ARMA model to predict the station demand to update the station status S ; then, the system implement dispatching strategy proposed in Section 4. More details are explained below.

Some preparations need to be done before the simulation as follows. The source data are formed as the Table 1 below. Firstly, we utilize the k-means algorithm to cluster all travel data according to its geographical location. Before that, some necessary work needs to be done, such as the conversion of coordinate coded via Geohash, data normalization, and data de-duplication. Then, we perform regional clustering via k-means, whose result is shown in Figure 5.

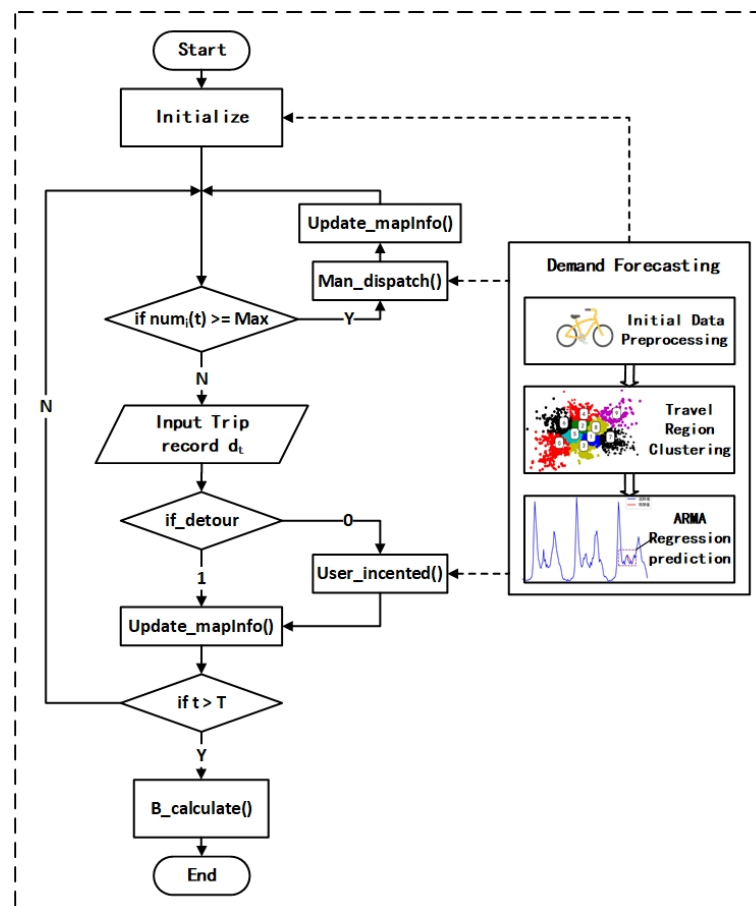


Figure 4. The procedure of hybrid dispatching scheme.

Table 1. Source data.

Orderid	Userid	Bikeid	Biketype	Start_time	Geohashed_start_loc	Geohashed_end_loc
163020	378959	22541	1	2017/5/10 0:00	wx4f5pf	wx4f5pe
3428602	140756	354271	1	2017/5/10 0:00	wx4g20b	wx4erbf
4580017	693933	459074	1	2017/5/10 0:00	wx4ersk	wx4ersf
4117589	482046	420604	1	2017/5/10 0:00	wx4dnv2	wx4dphv
2780838	1018307	293542	2	2017/5/10 0:00	wx4g2p1	wx4exbf

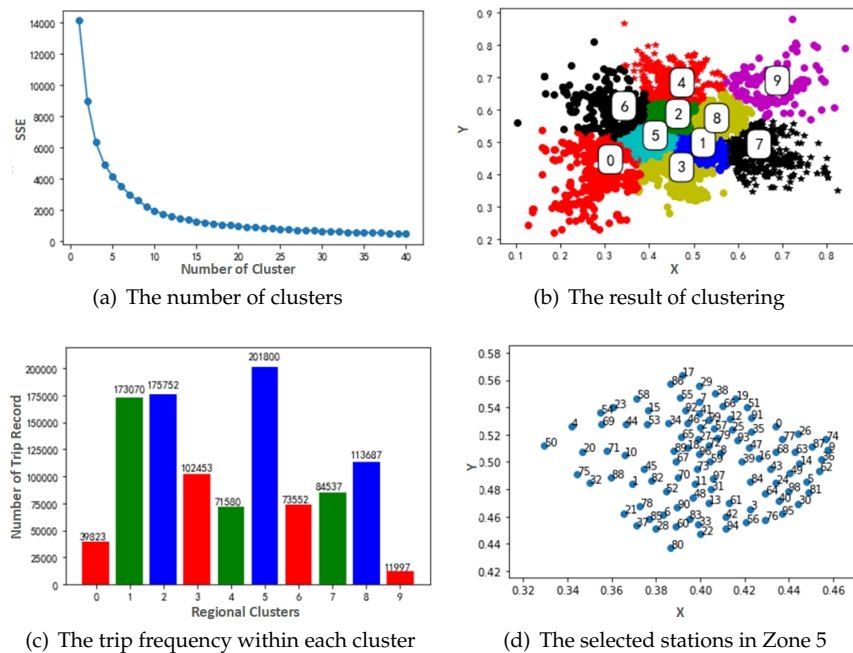


Figure 5. The division of regional clusters.

We calculate the Sum of Squared Error (SSE) of the cluster operation for a better number of clusters. As shown in Figure 5a, we set the number of clusters in the k-means algorithm to 10. The results of the clustering operation on the initial data are shown in Figure 5b. Then, we count the number of trip records in each region which are shown in Figure 5c. From the result, we can see that Zone 5 has the most trip records, which can be seen as a hot spot of trips and selected as the research area. In addition, we select 100 stations in Zone 5 and construct the undirected graph G , which is shown in Figure 5d.

Based on the filtered data in Zone 5, we remove some useless labels and resample it as the format like Table 2, where t is the current moment of simulation, s_type and e_type are the start and end of the trip record, and if_detour describes whether the user in trip record accepts detour for dispatching.

Table 2. Input data.

t	s_Type	e_Type	if_Detour
0	7	66	0
0	39	16	1
0	67	89	1
0	4	4	0
0	66	38	0

In addition to the trip record data, we also need to determine the status S of the station at the initial time, including the number of bicycles in all stations at the initial moment and the demand data at the next moment of all the stations, which can be defined as $S_i(0) = (num_i(0), pred_i(0))$. Several data sets of bicycle distribution conforming to Gaussian distribution are generated by Python to initial

the $num_i(0)$. The initial distribution of bicycles has a great impact on the performance of system dispatching, which usually acts like bicycles are concentrated in a few stations. In the paper, for a more comprehensive analysis of dispatching strategy, we generate multiple initial distribution data sets and utilize the standard deviation of 100 stations' initial distribution of bicycles σ to describe the concentration of bicycle distribution, which defined as Formula (6):

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=0}^{N-1} (num_i(0) - \overline{num(0)})^2} \quad (6)$$

Next, we need to predict the bicycle demand at each station in each time period. Based on the trip record shown in Table 2, we should first build the ARMA model for demand forecasting, which can be divided into three steps and illustrated as follows:

- (1) **Stationary correction of the trip data:** The premise of ARMA is that the series are stationary, which means that the statistical characteristics (mean, variance, correlation, etc.) will not change with different time Windows, and the mathematical expression is the same joint distribution. As shown in Figure 6a, we can determine that the trip record is a stationary sequence.
- (2) **Test of white noise:** Next, we will judge whether the original sequence is white noise or not. If it is white noise, then the sequence is random disturbance, which cannot be predicted and used. The sequence $\{e_0, e_1, e_2, \dots, e_t, \dots\}$ is white noise in three conditions as follows:

$$\begin{cases} E(e_t) = 0; \\ Var(e_t) = \sigma^2; \\ Cov(e_t, e_{t+h}) = 0, h \neq 0. \end{cases} \quad (7)$$

We utilize the Ljung–Box test to judge the white noise. By calculation, we get that $p = 1.47620 \times 10^{-17}$ when values are up to the lag of h . As a result, we can reject the original hypothesis and assume that the sequence is not a white noise sequence.

- (3) **Determination of time series lag:** Auto-Correlation Factor (ACF) and Partial Auto-Correlation Factor (PACF) are utilized to determine the order of the model. As shown in Figure 6b,c, the appropriate lags of p and q in ARMA model both are 2.

Finally, we build a discrete time dispatching simulation system via Python. As shown in Figure 4, dispatching system enters a state of cycle monitoring after the initializing of simulation. Next, the system perform manual dispatching and user incentive dispatching based on the status S and if_detour . After the timing is over, the system will collect the number of successful rental orders and the total revenue within the dispatching time. The following Table 3 provides specific simulation parameter settings.

Table 3. Simulation parameters' settings.

Parameter	Value
Duration $T(0.5 h)$	48
Threshold number of station Max	30
Remuneration of single trip $cost_t(RMB)$	2
Cost on single manual dispatching $cost_m(RMB)$	10
Cost on single user incentive dispatching $cost_u(RMB)$	1
Average trip distance of the shared bicycle $R(km)$	0.05
Maximum working distance of users $r(km)$	0.01
Maximum number of bicycles in manual dispatching k_0	50
The initial distribution of bicycles σ	7.45302, 8.56492
The percent of users receiving dispatching tasks	0, 0.2, 0.4, 0.6, 0.8, 1
The number of trip records	48365

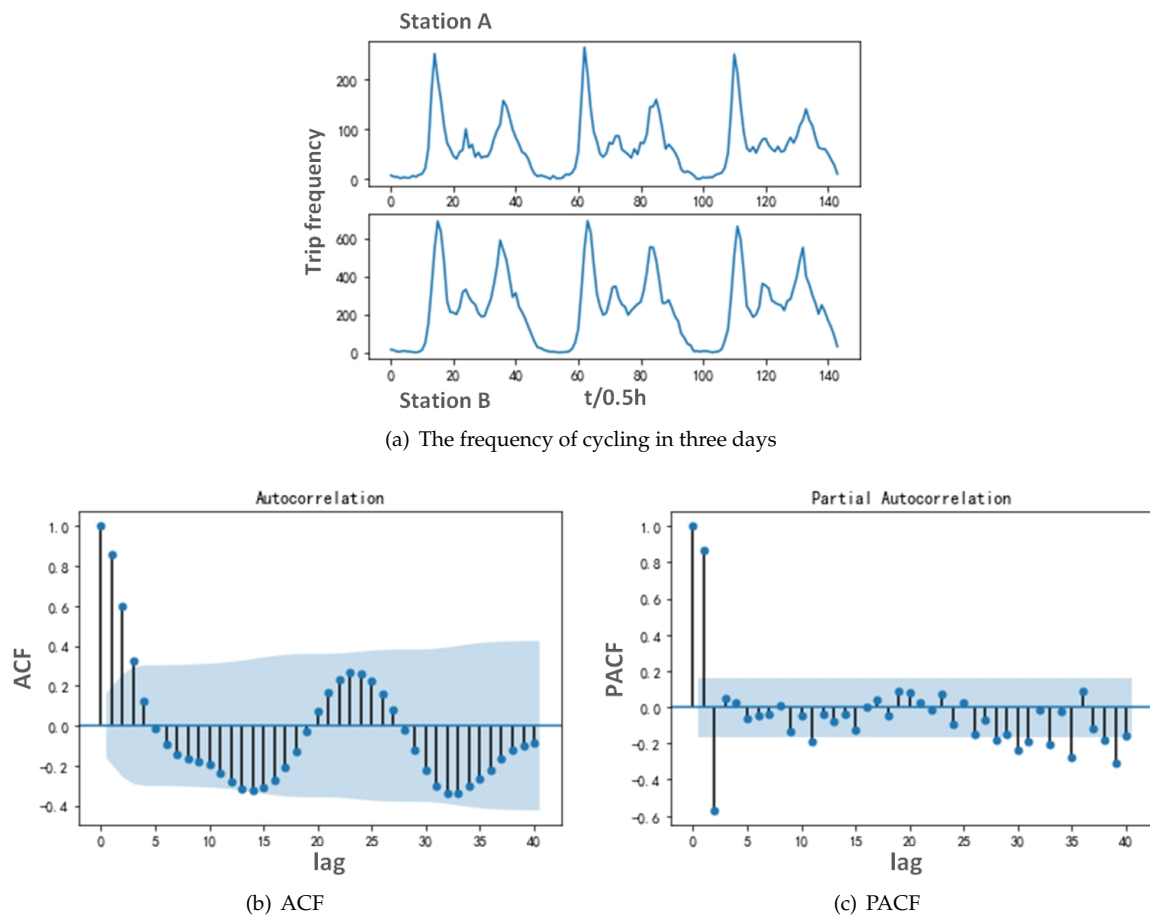


Figure 6. The determination of parameters in the ARMA model.

5.2. Results and Discussion

The results of the experiment will be analyzed in this section. Before simulation, we need to build a better ARMA model as described above. Below, we first analyze the results of the ARMA model.

We use the previously established ARMA model to perform curve fitting on the original data. As shown in Figure 7a, Residual Sum of Squares (RSS) is used to describe the performance of curve fitting, and our model has achieved good results. As an example, we select any period in the experimental data to predict the demand. After prediction and comparison with the original data, we got the results as shown in Figure 7b. We can see from the prediction results that the overall trends of the forecast curve and the true curve are similar, but the forecast is not accurate sometimes due to the small fluctuation of the bicycle rental, which can be ignored in subsequent dispatching schemes.

Next, we will perform an error analysis on the experimental results. In summary, the final results of the experiment will be affected by two factors which are the initial distribution of bicycles and the trip record. The following is a brief introduction.

Firstly, the initial distribution of bicycles on the map has a greater impact on experimental results. As shown in Formula (6), we generate two maps with different bicycle distributions, which can be indicated by parameter σ . In the traditional public bicycle system, a common dispatching mode is to dispatch excess bicycles at overflow stations directly to an underflow station, where the dispatching distance is taken as the optimization goal. We can call this type of dispatching scheme Commonly Used Manual Dispatching (CUMD). In addition, we refer to the manual dispatching strategy and the hybrid dispatching scheme including manual dispatching and user incentive dispatching proposed in this paper respectively as Manual Dispatching Proposed (MDP) and Hybrid Dispatching Proposed (HDP).

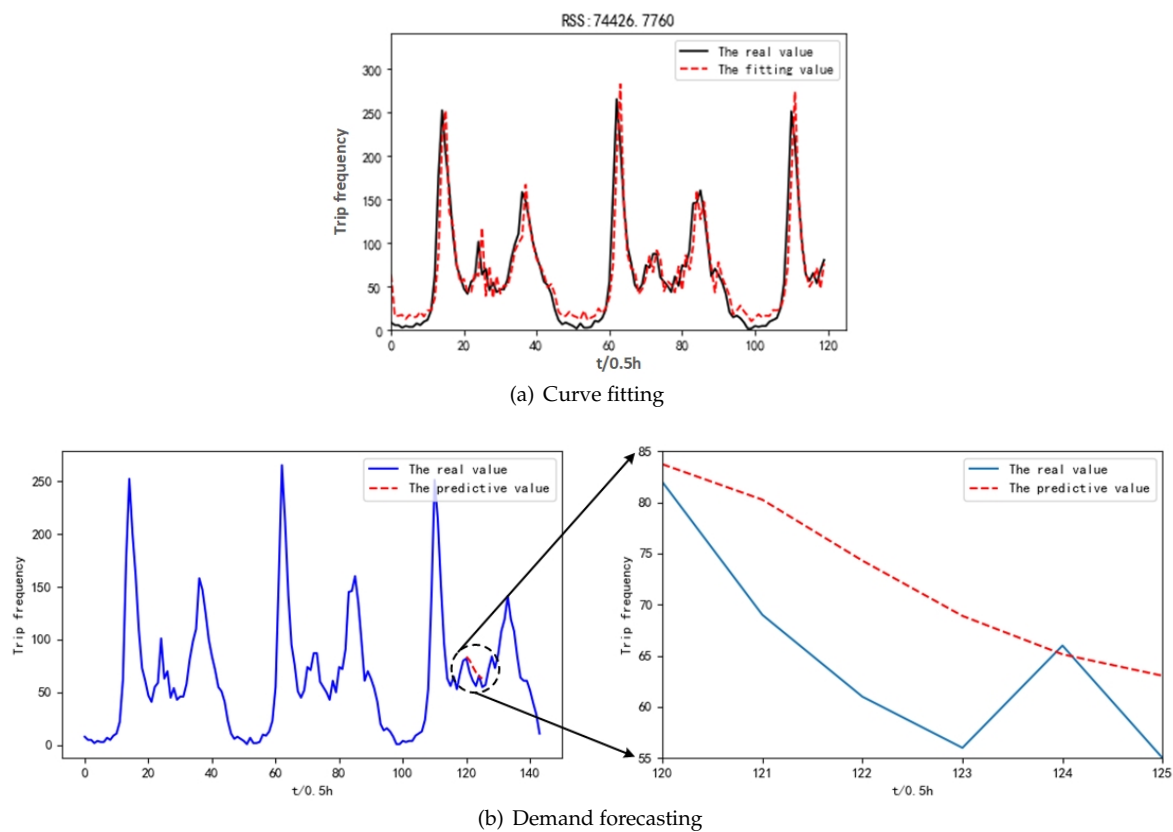


Figure 7. The result of predicting.

Secondly, the HDP proposed in this paper is also influenced by the proportion of riders in the travel record who are willing to accept user incentive dispatching. We set the step to 0.2 to constantly change the proportion of users in the trip record who are willing to accept dispatching tasks from 0 to 1, so as to conduct multiple sets of experiments.

Finally, the simulation results are as follows, which are listed to five significant digits. In Table 4, we can see the experimental results of user satisfaction for three dispatching methods. From the experimental results, we can see that **CUMD** is relatively behind **HDP** and **MDP** in terms of user satisfaction for almost 0.02. Since our experiment has more than 40,000 trip records, the improvement of user satisfaction exceeding 0.02 is a considerable improvement for the bicycle dispatching system. Furthermore, the **MDP** and **HDP** proposed in this paper are roughly the same in terms of user satisfaction. However, the value of HDP is going to change a little bit more due to the integration of user incentive dispatching and experimental results also prove it. As shown in Figure 8, the curves of the operator's total revenue B over time are shown. Firstly, the **HDP** and **MDP** schemes have almost the same performance in terms of operator revenue. Then, **HDP** is consistent with the other two strategies in the early stage of scheduling, but, after a period of scheduling, its performance in terms of operator revenue is better than the other two strategies.

Table 4. The satisfaction of user.

Dispatching Method	User's Satisfaction α
CUMD	0.767 ± 0.010
MDP	0.787 ± 0.002
HDP	0.788 ± 0.012

From the above simulation results, it shows that the dispatching scheme which combines professional dispatchers and ordinary users proposed in this paper is better than traditional manual dispatching to a certain extent. The manual dispatching strategy proposed in this article can accurately dispatch bicycles to some stations with greater demand according to the prediction results, so that more users can find a bicycle to rent. However, due to the higher cost of manual dispatching, we need to supplement it with a certain amount of user incentive dispatching, which combines the idea of crowdsourcing and assigns dispatching characters to ordinary users at a relatively small cost.

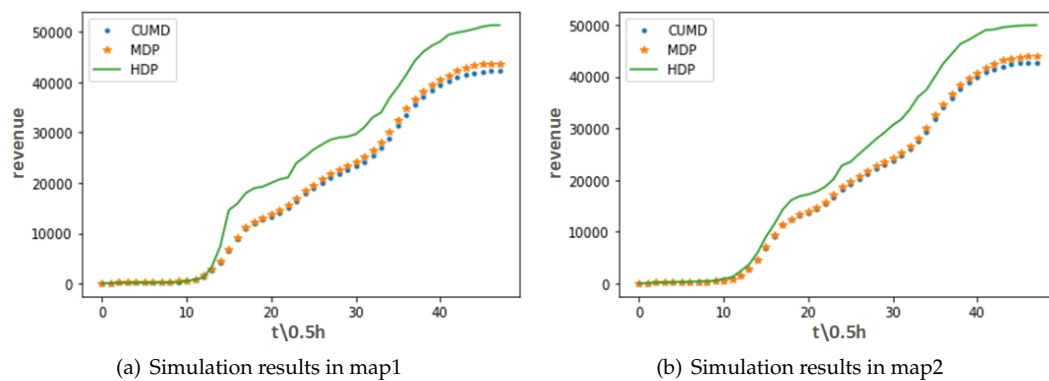


Figure 8. The revenue of operators.

6. Conclusions

In this paper, we suggest the idea of station imbalance $imb_i(t)$ to explain the difference at a time in the future between station demand and the actual number of bicycles at the station. Next, based on greedy ideas, using the sum of the imbalances of all stations as the objective function, we propose a hybrid scheduling strategy that includes manual scheduling and a user incentive schedule.

This paper proposed a hybrid dispatching strategy based on bicycle demand data to improve user riding satisfaction while reducing dispatching costs and improving overall revenue for bicycle operators. From the analysis of real-world traffic data in Shanghai from Mobike, it has been shown that the overall trend of daily bicycle trips during the week is cyclical, with a few peaks. Therefore, we clustered the bicycle records according to their starting point to extract bicycle trip stations, and then based on the historical data of bicycle trips, we built an ARMA model to predict the demand at potential stations in the short term. Then, we propose the concept of station imbalance $imb_i(t)$ to describe the gap between the demand for the station and the actual number of bicycles at the station at a moment in the future. Next, based on greedy ideas, using the sum of the imbalances of all stations as the objective function, we propose a hybrid dispatching strategy that includes manual dispatching and user incentive dispatching. Finally, we propose two indicators for evaluating dispatching strategies: user satisfaction and the total revenue of bicycle operators, and construct a discrete-time simulation model to simulate and analyze the hybrid scheduling strategy. The simulation experiment results show that the performance of the hybrid dispatching strategy proposed in this paper is better than the traditional bicycle dispatching strategy on both indicators, and it is less affected by the bicycle distribution in map, and can be adjusted to most dispatching scenarios.

For our future research, the forecast of bicycle demand is to be explored in depth with a balance of prediction speed and accuracy. The real-time dispatching strategy requires higher forecasting speed of the station demand forecasting method. Due to the irregular fluctuation of cycling trips caused by the emergency, the real-time scheduling strategy requires higher forecasting speed of the site demand forecasting method, and then improves the accuracy of the forecasting. Second, the incentive strategy in user incentive scheduling needs to be studied in depth, including the reasons that influence whether users accept crowdsourcing tasks, the formulation of appropriate incentive schemes, and the feasibility of user incentive strategies based on starting points. Last but not least, a better combination between

manual dispatching and user incentive dispatching should be studied, which can further reduce dispatching costs and improve users' satisfaction.

Author Contributions: Methodology, S.S., Z.W., and K.S.R.; Formal analysis and investigation, S.S. and Z.W.; Writing—original draft preparation, S.S., Z.W., and K.S.R.; Writing—Review and Editing, S.S., Z.W., and K.S.R.; Visualization, Z.W.; Supervision, L.S. and R.W.; Project Administration, S.S. and L.S.; Funding Acquisition, S.S., L.S., and R.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Science Foundation of P.R.China (Grant Nos. 61401221, 61873131, 61872196, 61701168, 61572261, and 61572260); China Postdoctoral Science Foundation (Grant No. 2016M601860); Postgraduate Research & Practice Innovation Program of Jiangsu Province (Grant Nos. SJCX19_0240, SJKY19_0823); NUPTSF (Grant No. NY220055); Teaching Reform from Project of Tongda College in NJUPT (Grant No. JG30618003).

Acknowledgments: The authors thank the financial support by all funding mentioned above and Shujaat Ali Zaidi's help for the proofreading of this paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Midgley, P. Bicycle-Sharing Schemes: Enhancing Sustainable Mobility in Urban Areas. Technical Report, Global Transport Knowledge Partnership & International Road Federation, 2011. Available online: <http://10.10.254.10/cache/3/03/www.un.org/6c48ce0251b0cfe3605d275f9709807e/Background-Paper8-P.Midgley-Bicycle.pdf> (accessed on 12 June 2018).
2. DeMaio, P. Bike-sharing: History, Impacts, Models of Provision, and Future. *J. Public Transp.* **2009**, *14*, 41–56.
3. DeMaio, P.; Gifford, J. Will Smart Bikes Succeed as Public Transportation in the United States? *J. Public Transp.* **2004**, *7*, 1–15.
4. Midgley, P. The Role of Smart Bike-sharing Systems in Urban Mobility. In *Journeys*; Red Avocado Film: Hochberg, Germany, **2009**; pp. 23–31.
5. Shen, S.; Wei, Z.Q.; Sun, L.J.; Su, Y.Q.; Wang, R.C.; Jiang, H.M. The Shared Bicycle and Its Network-Internet of Shared Bicycle (IoSB): A Review and Survey. *Sensors* **2018**, *18*, 2581.
6. OFO Bicycles. Available online: <http://m.ofo.so/> (accessed on 12 June 2018).
7. Mobike Bicycles. Available online: <https://mobike.com/cn/classic/> (accessed on 12 June 2018).
8. Sha, C.; Qiu, J.M.; Li, S.Y.; Qiang, M.Y.; Wang, R.C. A type of low-latency data gathering method with multi-sink for sensor networks. *Sensors* **2016**, *16*, 923.
9. Sha, C.; Zhong, L.H.; Bian, Y.; Song, D.D.; Ren, C.H. A type of energy-efficient target tracking approach based on grids in sensor networks. *Peer- Netw. Appl.* **2019**, *12*, 1041–1060.
10. Jefri, N.N.N.B.; Anuar, K.; Arjunan, S. Real Time Indoor Measurement of 2G, 3G and LTE Mobile Networks in Malaysia. In Proceedings of the 2016 IEEE 3rd International Symposium on Telecommunication Technologies (ISTT), Kuala Lumpur, Malaysia, 28–30 November 2016; pp. 19–24.
11. Raza, U.; Kulkarni, P.; Sooriyabandara, M. Low Power Wide Area Networks: An Overview. *IEEE Commun. Surv. Tutor.* **2017**, *19*, 855–873.
12. Ratasuk, R.; Mangalvedhe, N.; Zhang, Y.; Robert, M.; Koskinen, J.P. Overview of Narrowband IoT in LTE Rel-13. In Proceedings of the 2016 IEEE Conference on Standards for Communications and Networking (CSCN), Berlin, Germany, 31 October–2 November 2016.
13. Anand, S.; Regi, R. Remote Monitoring of Water Level in Industrial Storage Tanks Using NB-IoT. In Proceedings of the 2018 International Conference on Communication information and Computing Technology (ICCICT), Mumbai, India, 2–3 February 2018.
14. Shi, J.; Jin, L.; Li, J.; Fang, Z. A Smart Parking System Based on NB-IoT and Third-party Payment Platform. In Proceedings of the 2017 17th International Symposium on Communications and Information Technologies (ISCIT), Cairns, Australia, 25–27 September 2017.
15. Zhang, H.; Li, J.; Wen, B.; Xun, Y.; Liu, J. Connecting Intelligent Things in Smart Hospitals using NB-IoT. *IEEE Internet Things J.* **2018**, *5*, 1550–1560.
16. Chemla, D.; Meunier, Frédéric; Calvo, R.W. Bike sharing systems: Solving the static rebalancing problem. *Discret. Optim.* **2013**, *10*, 120–146.

17. Dell'Amico, M.; Hadjicostantinou, E.; Iori, M.; Novellani, S. The bike sharing rebalancing problem: Mathematical formulations and benchmark instances. *Omega* **2014**, *45*, 7–19.
18. Ho, S.C.; Szeto, W.Y. Solving a static repositioning problem in bike-sharing systems using iterated tabu search. *Transp. Res. Part E Logist. Transp. Rev.* **2014**, *69*, 180–198.
19. Li, Y.F.; Szeto, W.; Long, J.C.; Shui, C. A multiple type bike repositioning problem. *Transp. Res. Part B Methodol.* **2016**, *90*, 263–278.
20. Fu, S.Y.; Han, L.W.; Tian, Y.; Yang, G.S. Path planning for unmanned aerial vehicle based on genetic algorithm. In Proceedings of the 2012 IEEE 11th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC), Kyoto, Japan, 22–24 August 2012.
21. Martins, M.S.R.; Fuchs, S.C.; Pando, L.U.; Delgado, M.R. Pso with Path Relinking for Resource Allocation Using Simulation Optimization. *Comput. Ind. Eng.* **2013**, *65*, 322–330.
22. Horoba, C.; Sudholt, D. Running Time Analysis of ACO Systems for Shortest Path Problems. In *International Workshop on Engineering Stochastic Local Search Algorithms*; Springer: Berlin/Heidelberg, Germany, 2009.
23. Ahmed, S.A.A.; Safar, M.J.A. Shortest Path Planning for Noncircular Omnidirectional Mobile Robot. In *Intelligent Manufacturing & Mechatronics*; Springer: Singapore, 2018.
24. Gao, X.; Lee, G.M. Moment-based rental prediction for bicycle-sharing transportation systems using a hybrid genetic algorithm and machine learning. *Comput. Ind. Eng.* **2019**, *128*, 60–69.
25. El-Assi, W.; Salah Mahmoud, M.; Nurul Habib, K. Effects of built environment and weather on bike sharing demand: A station level analysis of commercial bike sharing in Toronto. *Transportation* **2017**, *44*, 589–613.
26. Feng, C.; Hillston, J.; Reijnders, D. Moment-based probabilistic prediction of bike availability for bike-sharing systems. In *International Conference on Quantitative Evaluation of Systems*; Springer: Cham, Switzerland, 2016.
27. Dickinson, E.; Euston, S.R. Statistical study of a concentrated dispersion of deformable particles modeled as an assembly of cyclic lattice chains. *Mol. Phys.* **1989**, *66*, 865–886.
28. García-Palomares, J.; Gutiérrez, J.; Latorre, M. Optimizing the location of stations in bike-sharing programs: A GIS approach. *Appl. Geogr.* **2012**, *35*, 235–246.
29. Dai, J.F.; Zhao, X.L.; Lin, J.X. Public bicycle system station deployment and scale prediction. *Adv. Mater. Res.* **2013**, *790*, 510–514.
30. Lin, F.; Zhang, L.; Xu, H.; Sun, Y. The application research of long-term and short-term prediction in the rental trends of public bicycle stations. *J. Comput. Theor. Nanosci.* **2016**, *13*, 9419–9425.
31. Bao, J.; He, T.; Ruan, S.; Li, Y.; Zheng, Y. Planning Bike Lanes based on Sharing-Bikes' Trajectories. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '17). Association for Computing Machinery, New York, NY, USA, 13–17 August 2017; pp. 1377–1386.
32. Evans, M. R.; Oliver, D.; Shekhar, S.; Harvey, F. Fast and exact network trajectory similarity computation: A case-study on bicycle corridor planning. In Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing (UrbComp '13). Association for Computing Machinery, New York, NY, USA, 11 August 2013.
33. Lv, M. X.; Liu, Y.N. Study of public bicycle system based on metro shuttle. *Urban Roads Bridges Flood Control* **2016**, *9*, 79–81.
34. 2017 Mobike Cup Algorithm Challenge on Biendata Competitions. Available online: <https://www.biendata.com/competition/mobike/> (accessed on 10 February 2020).

