Research on Scheduling Area Partition Method Based on Multiple Algorithms

Lefeng Li
College of Computer Science
Inner Mongolia University
Hohhot 010012, China
E-mail: 932164210@qq.com

Shanshan Wang
College of Computer Science
Inner Mongolia University
Hohhot 010012, China
Corresponding author e-mail: cswangss@imu.edu.cn

Abstract—**Shared bicycle, as a green means of transportation, is very popular among people and it is an important way for many people to travel daily. In recent years, with the increasing scale and frequency of bike sharing system, the unbalanced use of shared bicycle has a great impact on the users' experience, which is one of the main problems faced by current system operators. of Division of the traffic area can not only provide a new idea for solving the problem of unbalanced bike usage, but also provide a theoretical and practical basis for the planning, layout, construction, operation and scientific dispatching of shared bicycle system. However, there are few clear methods to study the partition method of shared dispatching area. To solve this problem, based on historical bicycle data, traffic station data, we analyze the rules of shared bicycle space-time characteristics and propose a method of dividing shared bicycle dispatching areas by combining K-medoids clustering, association rules and total demand constraint adjustment. We evaluate our approach on the New York City (NYC Citi Bike) bicycle sharing system and show the advantages of our approach for Large-scale station-level dispatching area optimization (beyond baseline approaches).**

Keywords-scheduling area partition method; bike sharing system; rebalance; clustering; Total demand optimization

I. INTRODUCTION

Bicycle sharing system is widely used domestically and abroad. It provides great help to solve the problem of "last kilometer" and traffic jam. A user can rent (i.e. check-out) a bike at a station near their origin and return (i.e. check-in) it to a station close to their destination. A record is generated when a bicycle is borrowed/returned, including the location of the origin station, the location of the destination station, and the duration of the ride.

However, bicycle sharing system still faces challenges in bike rebalancing between stations. Essentially, bike usage is constrained by time and location, so the traffic in the whole city will be unbalanced. For example, some stations may have a lot of bicycles returning, but some stations may not have available bicycles for users. In order to solve this problem, Most studies directly predict bicycle demand during a future period in order to avoid unbalanced problems and dispatchers manually schedule ahead of time, but the accuracy is not high,and when the number of stations is large, dispatchers need to schedule each station manually, which increases the workload and difficulty of delivery personnel and makes it difficult to achieve efficient and orderly scheduling. Thus, in order to reduce the workload and the difficulty of their jobs, therefor makes it difficult to achieve efficient and orderly scheduling.Thus,in order to reduce the workload and difficulty, we start with reducing the scale of the problem,which is to divide the whole traffic into different areas.The location, station-station trip frequencies as self-fluidity and the bike demands are taken into account,and the areas are adjusted by constraints to achieve the optimal results of the maximum balance within the areas. This not only reduces the workload and difficulty of dispatchers, but also experiments (see Chapter IV, Part C) show that the demand distribution of bike usage of cluster is more regular than that of station-level bike usage. Therefore, our proposed method lays a foundation for improving the accuracy of bike usage demand prediction in the future.

In these days, the scanning method and the clustering algorithm [1] are mainly used domestically and abroad to divide the vehicle scheduling area. Among them, the scanning method [2] is only applicable to the cases where the number of users is small and the distribution area is not large. Although the spatial clustering algorithm [3] applicable to the cases of a large number of customers and a large scheduling area, there is still a lack of criteria for reasonably allocating the weights of spatial and non-spatial attributes, or a lack of accurate methods to define the distance between feature attributes. If applied directly to dynamic scheduling of shared bicycle system, it may lead to large random errors.

Bike sharing system has a large number of stations and complex attributes, which makes it difficult to calibrate their attributes one by one. However, due to the self-fluidity of shared bicycle [4,5], there is a strong correlation between some stations.Therefore, if the association rules are used to collect stations with strong correlation and adopt K-medoids algorithm [6,7] and constraint adjustment for scheduling, it can effectively avoid large random errors.

Based on the above analysis, considering the actual bike demand of real-time dispatching of bike sharing system, we propose a scientific method for dividing the shared bicycle scheduling area by combining K-medoids clustering, association rules and restraint adjustment of total demand in the bike sharing system. In the first step of this method, K-medoids clustering is applied to the stations of bike sharing system.Then on the basis of the self-fluidity between stations and the transformation relationship between stations, the set of strong

DOI reference number:10.18293/SEKE2019-020.

association rules is screened out by using association rules. According to the constraints of the total demand in the areas, the total demand within the areas is optimized, and the regional division of dynamic dispatch of urban shared bicycle system is finally realized.

II. DESIGN OVERVIEW

We provide an overview of the symbols used in this paper (Table I) and problem definitions, as well as the description of design approach.

A. Basic Definitions

Definition 1: Station information. A station $S_{\mathrm{i}} = (id, lon_i, lat_i)$ denotes station information, where id represents the unique identity of each station, lon_i is the longitude of the station, lat_i is the latitude of the station.

Definition 2: Trip. A Trip $T_{\mathrm{r}} = (s_o, s_d, \tau_o, \tau_d)$ is a historical bike usage record, where s_o denotes the origin station, $\mathrm{s_d}$ is the destination station; τ_{o} and τ_{d} are the time when bike is checked out at s_o and checked in at $\mathrm{s_d}$,respectively.

Definition 3: Demand of bicycles. In time t, given a set of Check-out of station s_i , $\mathrm{O}_{S_i}(t) = \{O_{S_1}, O_{S_2}, ...O_{S_n}\}$ and check-in of station s_i , $\mathrm{I}_{S_i}(t) = \{I_{S_1}, I_{S_2}, ...I_{S_n}\}$,We want to get the demand of each station $S_i.d(t) = I_{S_i}(\mathrm{t}) - O_{S_i}(\mathrm{t})$.

*Problem Definition:*Scientific division of regions. Given a set of stations $S_{\mathrm{i}} = \{S_1, \ S_2, ...S_{\mathrm{n}}\}$, we want to cluster each station s_i to form $C_{1,\ \mathrm{i}} = \{C_{1,1}, C_{1,2}, ...C_{1,\ \mathrm{k}}\}$ clusters.

TABLE I. NOTATIONS

Notation	Description
N	Number of historical bike usage records
ρ	Coefficient of normal bike transaction records
n_b	Number of shared bikes
t	Days of data acquisition
λ	Coefficient of bike flow distribution
ω	Coefficient of on-frame mobile bike
$r_{i,j}$	Correlation coefficient from station i to station j
$n_{i,j}$	Number of bikes from station i to station j
n_i	Number of bikes flowing out from station i
S_i	The i^{th} station
T_r	A bike usage record
C_i	The i^{th} cluster
$O_{s_i}(t)$	Check-out of station S_i in time t
$I_{s_i}(t)$	Check-in of station S_i in time t

B. Design Methodology

Despite the time and location of the user's choice of borrowing is random, bikes are bound to be checked in at some station. Based on this simple observation, bike sharing system is decoupled in Figure 1 into two parts by analyzing the mobility [8,9] of bikes and characteristics.Based on historical bike usage records, we first use statistical methods to analyze the spatial and temporal distribution characteristics of bikes. Then, considering the space-time distribution characteristics of shared bikes, we propose a scientific method for dividing dispatching areas (see chapter III). Finally, the experimental results of NYC Citi Bike System show the advantages of our method.

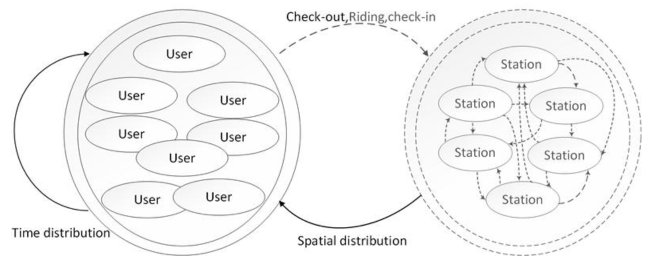

Figure 1. Components of a bike-sharing system

III. SCHEDULING AREA PARTITION ALGORITHMS

In this chapter, firstly, the spatial and temporal distribution characteristics of bike sharing system operation data are analyzed. The purpose is to grasp the operating regularity of the system, mine the trip patterns, obtain the macro operation rules of each station, and improve the quality of the dynamic scheduling of the bike sharing system.Then we put forward a scientific division method and show the specific steps of implementation through the above analysis. The goal of scientific partitioning is to transform the problem from a complex one (about 1,000 predictions per hour) to a simpler one, thereby reducing the complexity of the problem, making it easier to handle and helping to avoid over-fitting.

A. Spatiotemporal Analysis

Distribution Characteristics. As shown in Figure 2, from the macro-analysis [10] of the impact of month on the shared bicycle demand, it can be seen that there is a regular pattern of increasing demand from April to June. From June to September, the demand is stable but still in a high level. Selecting these months is conducive to dealing with the imbalance during the peak period. The figure on the right specifically shows the time distribution regular pattern of two different parameters, stations and time. It can be found that the early peak appears around 8 o'clock and the late peak appears around 5 o'clock.

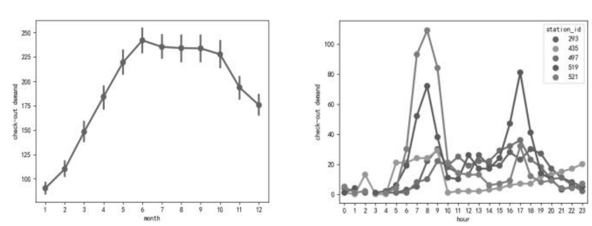

Figure 2. Law of time distribution

Analysis of Spatial Distribution Characteristics. Spatial distribution [11] characteristics analysis is based on each station as the research object. It analyses the distribution regularity of the whole city and the stations to which the vehicles borrowed from the station are returned or the vehicles returned to the station are borrowed from the station, and calculates the correlation between the stations. Station correlation refers to the flow correlation between two stations. We use the correlation coefficient to express their correlation. The larger the value, the more frequent the bicycle flows between the rental points, the greater the travel demand of users in this area. The formula can be expressed as follows.

$$\mathrm{r}_{i,j} = n_{i,j} * n_i^{-1} \tag{1}$$

Through the analysis of the spatial distribution characteristics of the stations, the borrow-return flow relationship between the stations in the system is determined, which provides the data basis for clustering the dispatching areas.

B. *Specific realization*

Figure 3 presents the iterative procedure of the partitioning method which organically combines three factors (location, self-fluidity and bike usage demand) of the stations.Stations within the same circle represent a cluster. The algorithm repeats the following three steps in each iteration: Geo-clustering, Strong Association Rule generation and constraint adjustment.

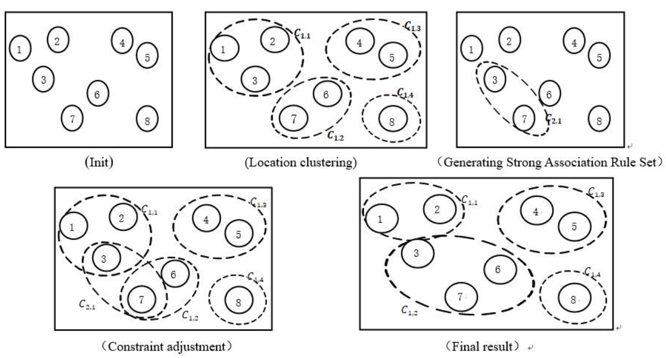

Figure 3. Partitioning method procedure

- *Location clustering.* According to the geographic location of each station, it is clustered into $K_1 = \{C_{1,k}\}_{k=1}^{k_1}$ by K-medoids method, this is the first time that location clustering is performed on all stations in the shared system".
- *Strong Association Rule Set generation.* check-in/ check-out between each station is calculated by the statistics of historical bike usage records, and the strong association rule set $\{C_{2,k}\}_{k=1}^{k_2}$ is screened by Aprior algorithm [12,13]. In this paper, we select 7:00-9:00 in the early peak period and 17:00-19:00 in the late peak period as the basis for screening strong association rule sets. In the process of preliminary classification of shared bicycle stations by association rules, the minimum support threshold $Sup_{\min}$ is very important, which determines the quality of clustering in the next step. If the value is too small, the correlation between the stations in the set is very weak, it will bring great errors to the later clustering to divide the dispatching area. If the value is too large, some stations with correlation may be screened out. When the next clustering division is carried out, most stations with less correlation will be introduced, which will also lead to larger errors in the result of division. Since there is no general method to determine the minimum support threshold, which is usually set artificially according to specific conditions, this paper considers that in a bike sharing system, the minimum support threshold should be determined according to various factors, and the expression is as follows.

$$Sup_{\min} = \frac{N * p}{t * \omega * n_b} * \lambda \tag{2}$$

A frequent itemset [14] is formed by selecting the records whose correlation coefficients are greater than those of the bike usage records. The relevant stations are put into the same set by using association rule algorithm. The stations in the set are the result of the users' free choice of the place to rent or return the bike when they travel. The principle of dividing the dispatching area is to excavate the travel rules of the users and balance the task of dispatching the vehicles. To improve scheduling efficiency, the correlation set meets the precondition of adjusting clustering in the next step.

- *Constraint adjustment.* The cluster $\{C_{2,k}\}_{k=1}^{k_2}$ obtained in step 2 and cluster $\{C_{1,k}\}_{k=1}^{k1}$ obtained in step 1 are calculated as follows.

$$\{C_{2,k}\}_{k=1}^{k_2} \cap \{C_{1,k}\}_{k=1}^{k_1} \tag{3}$$

where $\{C_{1,k}\}_{k=1}^{k_1}, \{C_{2,k}\}_{k=1}^{k_2} \neq \phi$ .If the result calculated by (3) is $C_i \{C_{1,k}\}_{k=1}^{k_1}$, it is the result of optimization, On the contrary, s_i in C_i is calculated in $\{C_{1,k}\}_{k=1}^{k_1}$ and $\{C_{2,k}\}_{k=1}^{k_2}$ as follows.

$$C_s(\mathrm{t}) = \min\left|\sum_{i=1}^{k} s_i . d(\mathrm{t})\right| \tag{4}$$

where the $C_S(\mathrm{t})$ value at $\{C_{1,k}\}_{k=1}^{k_1}$ is the smallest, then s_i is classified as $\{C_{1,k}\}_{k=1}^{k_1}$, and vice versa. Until all the collection in the k_1 clusters: $\{C_{2,k}\}_{k=1}^{k_2}$ are processed.

C. *Algorithm Complexity Analysis*

Time complexity [15,16]. The time complexity of our proposed method is mainly composed of k-medoids clustering and searching for the minimum total demand. In k-medoids algorithm, each point needs to be enumerated and the sum of its distances to all other points is obtained, so the complexity is $O(n^2)$. In addition,the time complexity of seeking the minimum total demand is $O(n^2)$ .To sum up, the time complexity of the whole algorithm is $O(n^2+n^2)$,i.e. $O(2n^2)$.

Spatial complexity[17]. The main memory overhead of the algorithm is the calculation of the cluster center and the total demand. The memory overhead can be effectively reduced by calculating the distance of a single data object at a time and the local demand in the morning and evening peak periods,which results in a spatial complexity of $O(n)$, so the spatial complexity of the whole algorithm is $O(n)$.

Algorithm 1: Partition Method Algorithm

Input: Station $\{S_i\}_{i=1}^{n}$, Trips$\{T_r\}_{i=1}^{n}$,iteration threshold K, parameters K_1>K_2

Output: $K_1 clusters\, C_{1,1}$, $C_{1,2}...C_{1,k_1}$

Cluster $\{S_i\}_{i=1}^{n}$ into K_1Clusters, $C_{1,1}$, $C_{1,2}...C_{1,k_1}$By K-medoids Based on locations;
Stations$\{S_i\}_{i=1}^{n}$ into K_2 Clusters, $C_{2,1}$, $C_{2,2}...C_{2,k_2}$By Aprior Algorithm;
Initialize k=1;
While k<K Do
  for a=1:n Do
    for i=1:n Do
      If C_i=$C_{2,a}\{S_i\}_{i=1}^{n} \cap C_{1,i}\{S_i\}_{i=1}^{n} \neq \emptyset$ and $C_i \neq C_{2,a}\{S_i\}_{i=1}^{n} \neq C_{1,i}\{S_i\}_{i=1}^{n}$ **Then**
        Calculate C_s(t) in the set of $C_{2,a}\{S_i\}_{i=1}^{n}$ and $C_{1,i}\{S_i\}_{i=1}^{n}$ respectively for S in C_i;
        If $S_i \in C_{1,i}\{S_i\}_{i=1}^{n}$, min ($C_s$(t) **Then**
          S_i is classified as cluster $C_{1,i}\{S_i\}_{i=1}^{n}$;
        Else $S_i$is classified as cluster $C_{2,a}\{S_i\}_{i=1}^{n}$;
  k=k+1;
Return K_1 $clusters$ $C_{1,1}$, $C_{1,2}...C_{1,k_1}$

Figure 4. An Algorithm of Partition Method

IV. EXPERIMENTS

In this section, we use our proposed method to construct a model for partitioning the scheduling area, and test our method on two data sets (station data and bike data [18]).

A. Data Collection

This paper uses two data sets, one is historical bike usage data set, the other is station data set. These data sets record data from April 1 to September 30, 2014. Through the statistical collation of the data set, the station information and the historical bike records are unified into bicycle data set, a total of 473,620 records were recorded. The detailed description below can be obtained in Table II.

TABLE II. DETAILS OF THE NEW YORK DATA COLLECTION IN 2014

Bike Data	**#Stations**	344
	#Bikes	6800
	#Records	5,359,995

B. Baseline & Metric

The method in our work to divide traffic dispatching areas is denoted as Partition Method（PM）by combining K-medoids clustering, association rules and total demand.In order to confirm the effectiveness of our algorithm, we carried out experiments to compare our method with the following baselines:

Bipartite Station Clustering(BSC)[19]. This method grouped individual stations into clusters according to their geographical locations and transition patterns. Finally, the whole traffic is divided into 23 groups.

Adaptive Capacity Constrained K-centers Clustering (CCKC)[20]. This method considers the distance between stations and the location of outliers, grouping outliers with other outliers, and setting up delivery personnel in outliers.

Metric. The metric we adopt to measure results are Sum Of The Squared Errors (SSE).

$$SSE=\sum_{i=1}^{k}\sum_{p\in C_i}\left|p-m_i\right|^2 \tag{5}$$

Where p is the sample point in C_i and m_i is the center of C_i .

C. Experimental result

Station-level partition method. Intuitively, the larger the number of clusters, the lower the prediction accuracy. When there is only one cluster, its usage demand is the whole traffic flow, which can be predicted accurately; when there are clusters, it means that each station forms a cluster, and the outflow/ inflow of the cluster fluctuates greatly, even if it can be predicted, but it is difficult to predict accurately. However, on the other hand, the number of clusters should not be too small, because if the cluster is too large, such as a cluster containing all stations, redistributing bicycles to the cluster cannot provide convenience for users. Therefore, we take the number of outliers as the baseline though many experiments, and finally the number of clusters is determined to be 23.

Similarly, we use another method to evaluate the effectiveness of our method.That's the elbow method SSE we talked about above. When the number of clusters K is less than the number of real clusters, the aggregation degree of each cluster will be greatly increased with the increase of K. When K reaches the real cluster number, the aggregation degree returns will be rapidly reduced with the increase of k, so the decrease of SSE will decrease sharply, and then become flat with the increase of K value. That is to say, the graph of the relationship between SSE and K is as follows: The shape of an elbow, and the K value corresponding to this elbow is the real clustering number of data. Obviously, As can be seen from Figure 5, when the SSE value is the smallest, the number of clusters K is still 23. Therefore, the effectiveness of our method is verified.

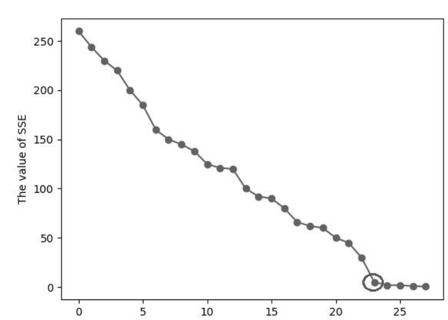

Figure 5. Evaluation of our method

Determination of Minimum Threshold. In addition, when we use Aprior algorithm to further filter the set of strong association rules, the minimum support threshold will be involved. In the process of borrowing and returning, because some records of data set are abnormal transaction records generated by the manual bicycle dispatching operation of the station, the coefficient of normal bicycle transaction records is taken as $\rho=0.98$; when selecting data samples, there are about 6800 bicycles in the bike sharing system, but a considerable number of bicycles are in the Off-Shelf state during the peak period, that is to say, they do not participate. With the flow of shared bicycles, the coefficient $\omega=0.9$ of mobile bicycles on the rack is taken. when shared bicycles are moving at the station, according to the results of data analysis, bicycles leased from one s may be returned to other 3-5 stations besides their own, so the distribution coefficient of bicycle flow is taken as $\lambda=0.25$, according to formula (2) ,the minimum support threshold $Sup_{\min}=0.1$ is obtained. The final result is obtained through the restraint adjustment of the total regional demand. (as shown in Figure 6.)

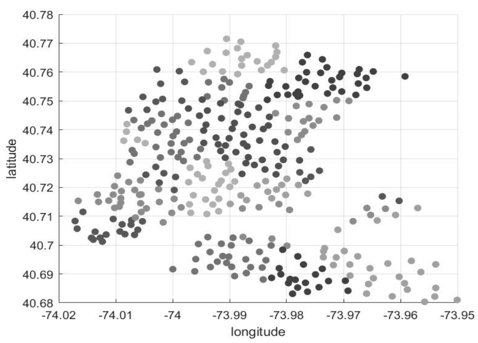

Figure 6. Clustering results

Performance comparison. In order to confirm the effectiveness of our model, we carried out experiments and compared our method with two baselines. CCKC (adaptive constrained central point clustering) and BSC (clustering based on K-means method according to the transformation relationship and geographical location between stations). The effectiveness and efficiency of the proposed PM are shown in Figure 7. It can be seen that for a given number of vehicles, we can concentrate on optimizing the stations and effectively find abnormal stations. With the increase of the number of vehicles scheduled, the number of outlier stations decreases rapidly. PM algorithm can help determine the minimum number of vehicles covering all target stations, or balance operation costs and the number of outlying stations.

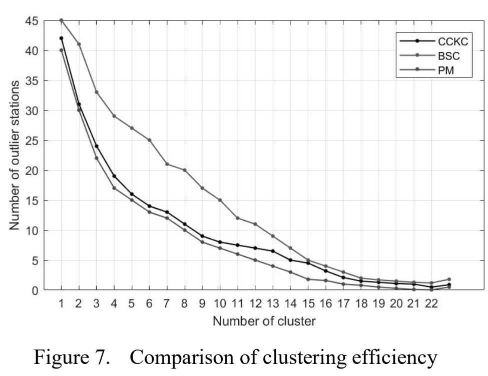

Figure 7. Comparison of clustering efficiency

Clustering Analysis of Shared Bicycle Usage Distribution. Figure 8 shows the demand of shared bicycle in different time at station level and class level. It can be concluded that the trend of shared bicycle demand is more stable after using the zoning algorithm. This demonstrates the effectiveness of our proposed partition method, and also provides the possibility for improving the prediction accuracy of shared bicycle demand in the traffic field.

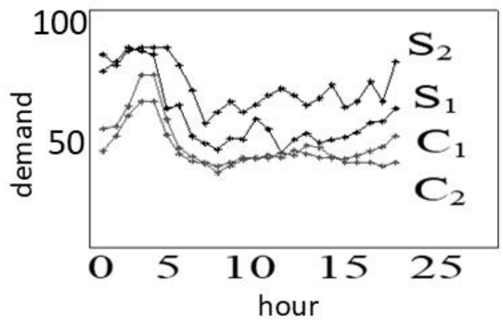

Figure 8. Station and cluster distribution

V. CONCLUSIONS

Based on the characteristics of complex relationship, large scale and large randomness among stations of bike sharing system, we consider the self-mobility of bikes among stations. We study urban bike sharing system with three algorithms: Association rules, K-medoids clustering and total demand restraint adjustment, and take New York City bike sharing system as an example to simulate and partition. This method takes into account the relationship between stations, the attributes of geographical location and the total demand for bicycles. Compared with CCKC and BSC methods, it is concluded that the number of outliers in this method is more stable and the value of SSE is the smallest. Furthermore, The area obtained by our algorithm is more stable than that of single station, which also shows that our method provides a theoretical basis for improving the accuracy of traffic flow prediction.

ACKNOWLEDGMENT

This work is supported by Programs of National Natural Science Foundation of China (No:71461023) .

REFERENCES

[1] Saberi M , Ghamami M , Gu Y , et al. Understanding the impacts of a public transit disruption on bicycle sharing mobility patterns: A case of Tube strike in London[J]. Journal of Transport Geography, 2018, 66:154-166.

[2] Li Y , Zheng Y , Zhang H , et al. Traffic prediction in a bike-sharing system[C]// the 23rd SIGSPATIAL International Conference. ACM, 2015.

[3] Liu J , Sun L , Chen W , et al. Rebalancing Bike Sharing Systems: A Multi-source Data Smart Optimization[C]// Acm Sigkdd International Conference on Knowledge Discovery & Data Mining. ACM, 2016.

[4] Yang Z , Hu J , Shu Y , et al. Mobility Modeling and Prediction in Bike-Sharing Systems[J]. 2016.

[5] Li H , Hong L Y , Mo Y C , et al. Restructuring performance prediction with a rebalanced and clustered support vector machine[J]. Journal of Forecasting, 2018(4).

[6] J. Liu, Q. Li, M. Qu, W. Chen, J. Yang, X. Hui,H. Zhong, and Y. Fu. Station site optimization in bike sharing systems. In ICDM 2015. IEEE, 2015.

[7] Holmgren J , Moltubakk G , Jody O'Neill. Regression-based evaluation of bicycle flow trend estimates[J]. Procedia Computer Science, 2018, 130:518-525.

[8] Dell'Amico, Mauro, Iori M , Novellani S , et al. A destroy and repair algorithm for the Bike sharing Rebalancing Problem.[J]. Computers & Operations Research, 2016, 71(C):149-162.

[9] Zhang J , Pan X , Li M , et al. Bicycle-Sharing System Analysis and Trip Prediction[J]. 2016.

[10] Longbiao Chen, Daqing Zhang, Leye Wang, Dingqi Yang, Xiaojuan Ma, Shijian Li, Zhaohui Wu, Gang Pan, Thi-Mai-Trang Nguyen, and Jérémie Jakubowicz.2016. Dynamic cluster-based over-demand prediction in bike sharing systems.In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing. ACM, 841–852.

[11] Mike Benchimol, Pascal Benchimol, Benoît Chappert, Arnaud De La Taille, Fabien Laroche, Frédéric Meunier, and Ludovic Robinet. 2011. Balancing the stations of a self service "bike hire" system. RAIRO-Operations Research 45, 1 (2011), 37–61.

[12] Gerardo Berbeglia, Jean-François Cordeau, Irina Gribkovskaia, and Gilbert Laporte. 2007. Static pickup and delivery problems: A classification scheme and survey. TOP 15, 1 (2007), 1–31.

[13] Lin L , He Z , Peeta S . Predicting Station-level Hourly Demands in a Large-scale Bike-sharing Network: A Graph Convolutional Neural Network Approach[J]. 2017.

[14] Faghih-Imani A , Hampshire R , Marla L , et al. An empirical analysis of bike sharing usage and rebalancing: Evidence from Barcelona and Seville[J]. Transportation Research Part A: Policy and Practice, 2017, 97:177-191.

[15] Zhang L , Tang S , Yang Z , et al. Demo: Data Analysis and Visualization in Bike-Sharing Systems[C]// International Conference on Mobile Systems. ACM, 2016.

[16] Oliveira G N , Sotomayor J L , Torchelsen R P , et al. Visual Analysis of Bike-Sharing Systems[J]. Computers & Graphics, 2016, 60:119-129.

[17] Bouveyron C , Côme, Etienne, Jacques J . The discriminative functional mixture model for a comparative analysis of bike sharing systems[J]. Dissertations & Theses - Gradworks, 2016, 9(4).

[18] http://www.citibikenyc.com/system-data.

[19] Chen T , Lu S . Robust Vehicle Detection and Viewpoint Estimation with Soft Discriminative Mixture Model[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2016:1-1.

[20] Li Q L , Chen C , Fan R N , et al. Queueing Analysis of a Large-Scale Bike Sharing System through Mean-Field Theory[J]. 2016.

[21] Tomaras D , Boutsis I , Kalogeraki V . Lessons Learnt from the analysis of a bike sharing system[C]// International Conference. 2017.

[22] O'Mahony E , Shmoys D B . Data Analysis and Optimization for (Citi)Bike Sharing[C]// Twenty-ninth Aaai Conference on Artificial Intelligence. AAAI Press, 2015.

[23] Liu J , Li Q , Qu M , et al. Station Site Optimization in Bike Sharing Systems[C]// 2015 IEEE International Conference on Data Mining (ICDM). IEEE Computer Society, 2015.

[24] Basch C H , Ethan D , Zybert P , et al. Public Bike Sharing in New York City: Helmet Use Behavior Patterns at 25 Citi Bike™ Stations[J]. Journal of Community Health, 2014, 40(3):530-3.

[25] Chen Q , Sun T . A model for the layout of bike stations in public bike-sharing systems[J]. Journal of Advanced Transportation, 2015, 49(8):884-900.

[26] Vogel P , Saavedra B A N , Mattfeld D C . A Hybrid Metaheuristic to Solve the Resource Allocation Problem in Bike Sharing Systems[J]. 2014.

[27] Li Z , Zhang J , Gan J , et al. Large-Scale Trip Planning for Bike-Sharing Systems[C]// 2017 IEEE 14th International Conference on Mobile Ad Hoc and Sensor Systems (MASS). IEEE Computer Society, 2017.

[28] Bao J , Xu C , Liu P , et al. Exploring Bikesharing Travel Patterns and Trip Purposes Using Smart Card Data and Online Point of Interests[J]. Networks and Spatial Economics, 2017, 17(4):1231-1253.

[29] Bordagaray M , dell'Olio, Luigi, Fonzone A , et al. Capturing the conditions that introduce systematic variation in bike-sharing travel behavior using data mining techniques[J]. Transportation Research Part C: Emerging Technologies, 2016, 71:231-248.

[30] Chen L , Xiaojuan M A , Nguyen T M , et al. Understanding bike trip patterns leveraging bike sharing system open data[J]. Frontiers of Computer Science, 2017, 11(1).

[31] Cagliero L , Cerquitelli T , Chiusano S , et al. Predicting critical conditions in bicycle sharing systems[J]. Computing, 2017, 99(1):39-57.

[32] Zhang J , Yu P S . Trip Route Planning for Bicycle-Sharing Systems[C]// 2016 IEEE 2nd International Conference on Collaboration and Internet Computing (CIC). IEEE, 2016.

[33] Faghih-Imani A , Anowar S , Miller E J , et al. Hail a cab or ride a bike? A travel time comparison of taxi and bicycle-sharing systems in New York City[J]. Transportation Research Part A Policy and Practice, 2017, 101:11-21.

[34] Zhang J , Pan X , Li M , et al. Bicycle-Sharing System Analysis and Trip Prediction[J]. 2016.

[35] Benarbia T, Labadi K, Darcherif A M, et al. Real-time inventory control and rebalancing in bike-sharing systems by using a stochastic Petri net model[C]// International Conference on Systems & Control. 2018.