The rebalancing of bike-sharing system under flow-type task window

Zihao Tian[a], Jing Zhou[a,*] , W.Y. Szeto[b], Lixin Tian[c,*], Wenbin Zhang[d]

[a] School of Management and Engineering, Nanjing University, Nanjing 210093, China

[b] Department of Civil Engineering, The University of Hong Kong, Hong Kong

[c] School of Mathematical Sciences, Nanjing Normal University, Nanjing, Jiangsu 210023, China

[d] Department of Mathematical Science, Nanjing University of Science & Technology，Taizhou Institute of Science & Technology, Taizhou 225300, Jiangsu, China

Abstract: With the growing importance of bike-sharing systems, this paper designs a new framework to solve rebalancing problem. It contains two aspects: dynamic rebalancing within each station and static rebalancing among stations. Firstly, we give a new flow-type task window (F-window) by defining the consistency index of travelers. It is more suitable as a task window for rebalancing than time-type task window (T-window) based on three aspects analysis. Through three assumptions, the temporal-distribution learning model including task window and station storage configuration, are built to realize new dynamic rebalancing. The spatial-distribution learning method is introduced to divide management areas for static rebalancing. The empirical results show that F-window can better match the strong time-sensitive of demand fluctuation. Compared with traditional rebalancing needs hours, each rebalancing within a station can be completed within average 4 minutes. By setting the station storage configuration, it makes rebalancing in this paper meets the demand of 28.3 times the hourly rebalancing within one week. And the number of vehicles visiting stations has dropped below 20%.

Keywords: Dynamic rebalancing, Static rebalancing, Spatial-temporal distribution learning, Bike-sharing system, Inventory threshold, Flow-type task window, Community detection

1. Introduction

Rebalancing problem is a hot issue in the research of bike-sharing system. Because travelers usually have short distance one-way trips, the time and space distribution of travelers' demand is extremely uneven, and often there are cases where the station has no bike to borrow or no dock to return (Raviv and Kolka, 2013, Cyrille et al.,2016, Ahmadreza et al., 2017, Behzad et al., 2020). Therefore, one method performed by operators is known as a Pickup and Delivery Problem (PDP) (Erdoğan et al., 2015, Li et al., 2016, Legros, 2019). Moreover, with the bike-sharing system developing quickly, the models and algorithms also need to be improved to solve large-scale optimization problem (Chemla et al., 2013a, Shui and Szeto, 2014, Han et al., 2015).

Depending on the way of rebalancing, it can be modeled by a dynamic or a static approach (Chemla et al., 2013b). In the literature, it is more common to find researchers dealing with a static approach. Static rebalancing usually takes place at night and based on the initial distribution of bikes or the expected-demand for the next day (Kadri et al., 2016). According to the objective function, it can be mainly divided into two aspects: some researchers work on minimization of travel cost or distance (Benchimol et al., 2011, Chemla et al., 2013, Dell'Amico et al., 2014, Erdoğan et al., 2014, Lin and Chou, 2012, Raidl et al., 2013, Rainer et al., 2015, Schuijbrock et al., 2017). Others focus on

* Corresponding authors. E-mail address: jzhou@nju.edu.cn(J.Z.); tianlx@ujs.edu.cn(L.T.)

the minimization of user dissatisfaction or its variants (Raviv et al., 2013, Angeloudis et al., 2014, Szeto et al., 2016, Ho and Szeto, 2017). Pal and Zhang (2017) explored the static management method of free-floating bike-sharing systems.

As bike-sharing systems have developed in many large cities, solving large-scale dynamic rebalancing problems with limited time is still a difficult problem. Although dynamic rebalancing increases the frequency of vehicles visiting stations (Contrado et al., 2012, Kloimullner et al., 2014), there is still a lot of demand lost. This is because the speed of dynamic rebalancing cannot match the strong time-sensitive demand fluctuation of travelers. (Zhang et al., 2017; Shui and Szeto, 2018). The results of dynamic rebalancing mainly focus on the following categories: minimum travel costs (Contrado et al., 2012, You, 2019, Brinkmann et al., 2019), minimization of user dissatisfaction and travel costs (Zhang et al., 2017, Shui and Szeto, 2018) or minimizing deviation from the target fill levels of each station (Kloimullner et al., 2014, Brinkmann et al., 2016). Pfrommer (2014) and Zhang et al., (2019) studied the price strategy in dynamic rebalancing. Leonardo et al., (2018) summarized the dynamic management framework of bike-sharing system. Caggiani et al., (2018) and Du et al., (2019) also built a management framework for free-floating bike-sharing systems.

Operators hope that rebalancing will enable the supply to match the demand in real time. This requires further data-driven exploration of the changing rules of demand. Some researchers used different methods to explore travelers' behavior (Zhao et al. 2015, Maria et al. 2016, Wei et al., 2019). But in the above studies, demand changes characteristics in temporal and spatial distribution are less combined in rebalancing. So the following two problems have become the trend of rebalancing research.

Question 1. How to build a dynamic model to match demand with strong time sensitive by using space for time, which means to transform the real-time rebalancing problem among stations into a more convenient study within each station, and obtain a station inventory configuration and a F-window division for rebalancing to reduce the loss of demand during peak hours.

Question 2. Facing with large-scale bike-sharing systems, how to divide different management areas which are more useful than simple administrative division to achieve an overall static rebalancing of bike-sharing system under limited time.

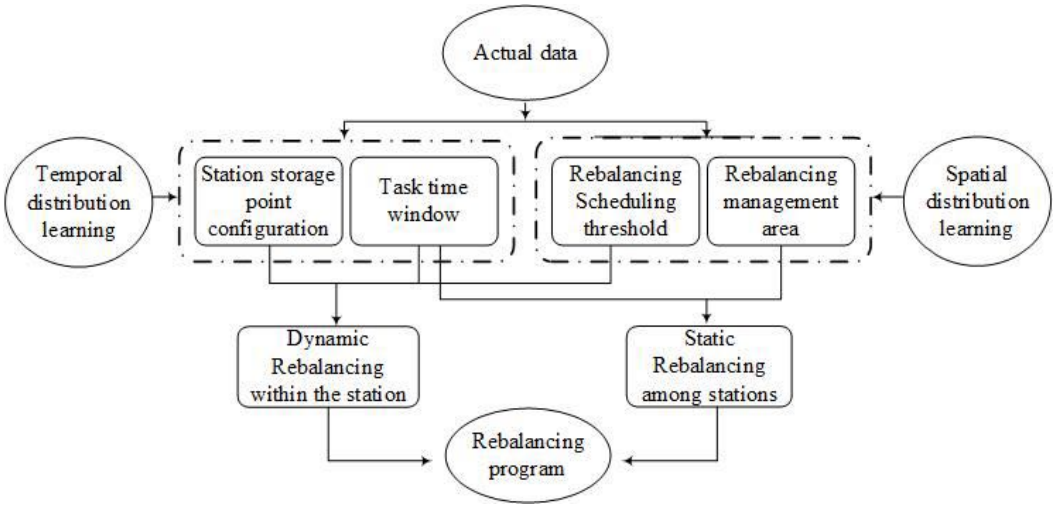


Fig.1 New bike rebalancing framework

This paper focuses on the above questions by temporal network theory and provides a new method for data mining (Banerjee et al., 2013, Yang et al., 2015). Based on three assumptions about space for time, rebalancing time within a station and the inventory recovery rate, we define the

consistency index that reflects demand fluctuation law through temporal distribution learning, give a new flow-type task window division by using this index and obtain a nonlinear time series related to time intervals under this division. Then, the initial capacity of each station storage point is determined under flow-type task window to form the temporal distribution learning of spatial instead of time and realize dynamic rebalancing within each station. Furthermore, the community detection method and robustness theory are used to determine the management area of system and starting threshold of station rebalancing, and the spatial distribution of bicycle data is learned to build management area of bike-sharing system. The new rebalancing scheme consists of two parts: dynamic rebalancing within each station and static rebalancing among the stations (See, Fig.1).

The dynamic rebalancing part is different from previous literatures (See, Table 1), our main contributions are: a) the task window is divided according to equal flow based on the spatial-temporal distribution of demands. The time interval under each equal-flow task window is determined by the travel direction, and the interval length is nonlinear and unfixed; b) we use the robustness theory of complex networks to calculate the starting threshold of each station. c) through the assumption of setting storage point in stations, rebalancing among stations is converted into rebalancing within each station. It can better match the strong time-sensitive of demand fluctuation; d) the scale of bike-sharing system reaches 1089 stations.

Table 1 Features of existing dynamic rebalancing scheme

	Window type	Window length	Flow per window	Threshold computing	Rebalancing type	Number of stations
Contardo et al., 2012	*T*	*C*	*V*	*N*	*AMS*	100
Kloimullner et al., 2014	*T*	*C*	*V*	*N*	*AMS*	90
Pfrommer et al., 2014	*T*	*C*	*V*	*N*	*AMS*	354
Regue and Recker, 2014	*T*	*C*	*V*	*Y*	*AMS*	61
Zhang et al., 2017	*T*	*C*	*V*	*N*	*AMS*	200
Chiariotti et al., 2018	*T*	*C*	*V*	*N*	*AMS*	280
Shui and Szeto, 2018	*T*	*C*	*V*	*N*	*AMS*	180
Brinkkmann et al., 2019	*T*	*C*	*V*	*N*	*AMS*	169
You, 2019	*T*	*C*	*V*	*N*	*AMS*	40
Warrington and Ruchti, 2019	*T*	*C*	*V*	*N*	*AMS*	102
This study	*F*	*V*	*C*	*Y*	*WIS*	1089

In Table 1, *T* stands for the task window of which the time interval is fixed but the flow is unfixed, called as T-window. *F* stands for the task window of which the flow is fixed but the time interval is unfixed, called as F-window. *N* stands for No, *Y* for Yes, *C* for constant, *V* for variety, *AMS* for rebalancing among stations, and *WIS* for rebalancing within each station.

The static rebalancing part is different from the previous literatures (See, Table 2), and our main distributions are: a) Due to the existence of storage warehouses (See, Assumption 1) within each station, The total daily demand for some stations can be huge, reaching up to 255 per day (See, Fig. 17). Therefore, we need to remove the limit on the number of each vehicle visits to the same station; b) We use the community detection method to divide total stations into five management clusters in order to deal with large-scale static rebalancing (more than 1000 stations). We also collected the actual road network information in order to make the results more accurate.

Compared with the scheme based on hourly dynamic rebalancing, we find that the scheme in this paper can match the change of system strong time-sensitive demand more effectively. Compared with traditional rebalancing needs hours, each rebalancing within a station can be completed within average 4 minutes. The cumulative station satisfaction frequency and level of demand satisfaction far exceed the hourly rebalancing scheme 31.82 and 28.31 times respectively, greatly reducing the demand loss. The number of stations visited and the total amount of bikes during vehicle rebalancing are also less than the hourly rebalancing scheme, reaching only 19.31% and 7.84% of the hourly dynamic rebalancing scheme, which effectively alleviates travel cost of vehicles. The structure of this paper is as follows: Section 2 gives the definition, assumption and rebalancing modeling. Section 3 is the empirical evidence of the rebalancing approach. Section 4 is the conclusion.

Table 2. Features of Existing Static Rebalancing

	Administrative division	Management area	Data type	Station size	$max\{\Delta Q_i\}$	Distance Type
Chemla et al., 2013	*Y*	*N*	*S*	100	30	
Erdoğan et al., 2014	*Y*	*N*	*S*	50		
Han et al., 2015	*Y*	*N*	*S*	1000		
Kadri et al., 2016	*Y*	*N*	*S*	100	10	
Li et al., 2016	*Y*	*N*	*S*	180		
Lin et al., 2012	*Y*	*N*	*R*	30		*A*
Angeloudis et al., 2014	*Y*	*N*	*R*	30	28	
Dell'Amico et al., 2014	*Y*	*N*	*R*	116	20	
Forma et al., 2015	*Y*	*Y*	*R*	200		
Schuijbroek et al., 2017	*Y*	*N*	*R*	135	50	*E*
Ho and Szeto, 2017	*Y*	*N*	*R*	518		
This study	*N*	*Y*	*R*	1089	255	*A*

In Table 2, *N* stands for No, *Y* for Yes; *S* represents simulation data, *R* represents real data; *A* means the actual distance among stations, *E* represents the Euclidean distance among stations; ΔQ_i represents the number of demand in station V_i in static rebalancing.

2. A new method of rebalancing: learning rebalancing

2.1. Definitions and Assumptions

2.1.1 Public bike-sharing network

This paper gives the following definitions and assumptions in order to establish the new rebalancing scheme of the bike-sharing system.

Definition 1. (1) We build the borrowing bike-sharing network (B-BN) by using stations as nodes and borrowing records as directed edges. (2) We build the returning bike-sharing network (R-BN) by using stations as nodes and returning records as directed edges.

Temporal window $\left\{\left[t_\alpha, t_\alpha + \Delta t_\alpha\right]\right\}_{\alpha=1}^{N}$ of temporal network $\left\{G^\alpha\left(V,E\right)\right\}_{\alpha=1}^{N}$ can well describe the time-varying characteristics of travelers, where N represents the number of temporal windows, and α is the order number. In temporal window $\left\{\left[t_\alpha, t_\alpha + \Delta t_\alpha\right]\right\}_{\alpha=1}^{N}$, the temporal networks for B-BN and R-BN are recorded separately $\left\{G_B^\alpha\left(V,E\right)\right\}_{\alpha=1}^{N}$ and $\left\{G_R^\alpha\left(V,E\right)\right\}_{\alpha=1}^{N}$, where V represents the set of stations, and E represents a set of directed paths between stations.

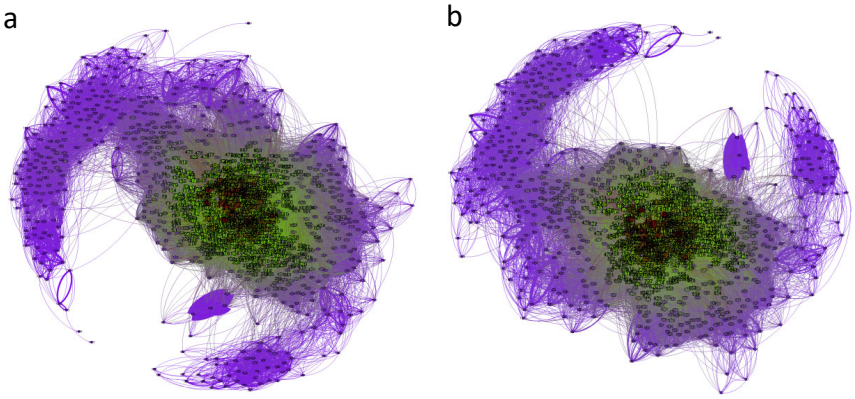


Fig. 2 The network structure of the bike-sharing network of Nanjing City (a) B-BN; (b) R-BN.

From definition 1, it can be seen that B-BN and R-BN are both edge-directed networks, and the weights of each edge representing the borrowing or returning times are different (See, Fig. 2). For temporal network $\left\{G_B^\alpha\left(V,E\right)\right\}_{\alpha=1}^{N}$, the out-degree of each node describes the demand from each station. If several bikes from station V_i go to the same station V_j within one temporal window $\left[t_\alpha, t_\alpha + \Delta t_\alpha\right]$, we consider that these travelers' direction from station V_i are consistent. So the out-degree $k^{\alpha,out}$ corresponding to the station V_i in temporal network $G_B^\alpha\left(V,E\right)$ is equal to 1. Next we will introduce the concept of consistency index in order to describe the law of demand fluctuation in each temporal window.

2.1.2 Task window

Definition 2. We use the probability $P\left(1^{\alpha,out}\right)$ of node occurring out-degree $k^{\alpha,out}=1$ in temporal window $\left[t_\alpha, t_\alpha + \Delta t_\alpha\right]$ of the $G_B^\alpha\left(V,E\right)$ as the consistency index.

Definition 3. Set $\left\{\left[t_\alpha, t_\alpha + \Delta t_\alpha\right]\right\}_{\alpha=1}^{N}$ as an arbitrary non-overlapping time window on $\left\{G_B^\alpha\left(V,E\right)\right\}_{\alpha=1}^{N}$ of temporal network for the borrowing bikes. If the consistency index $P\left(1^{\alpha,Out}\right)$ in temporal window network $G_B^\alpha\left(V,E\right)$ is satisfied:

$$\frac{1}{N}\sum_{\alpha=1}^{N} P\left(1^{\alpha,out}\right) \geq c_1, \text{ and } \min_{\alpha} P\left(1^{\alpha,out}\right) \geq c_2 \qquad (1)$$

We regard $\left\{\left[t_\alpha, t_\alpha+\Delta t_\alpha\right]\right\}_{\alpha=1}^{N}$ as effective partition based on the consistency index $P\left(1^{\alpha,out}\right)$ under the parameter c_1 and c_2 in the temporal network $\left\{G_B^\alpha\left(V,E\right)\right\}_{\alpha=1}^{N}$, where c_1 and c_2 are called the consistency factor.

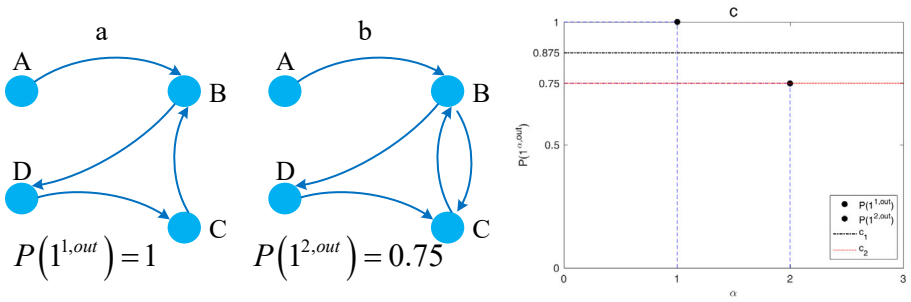


Fig. 3 Examples of consistency indices and consistency factors under $\left\{\left[t_\alpha, t_\alpha+\Delta t_\alpha\right]\right\}_{\alpha=1}^{2}$

For example, let $\left\{\left[t_\alpha, t_\alpha+\Delta t_\alpha\right]\right\}_{\alpha=1}^{2}$ be the non-overlapping windows of bike-sharing system network $\left\{G^\alpha\left(V,E\right)\right\}_{\alpha=1}^{2}$. In fig 3 (a) and (b), each node represents a station. Traveler's riding records constitutes edges. In fig. 3 (a), the out-degree of each node under the window $\left[t_1, t_1+\Delta t_1\right]$ is 1, according to definition 2, $P\left(1^{1,out}\right)=1$. In fig. 3 (b), out-degree of node A, C and D is 1 under the time window $\left[t_2, t_2+\Delta t_2\right]$ respectively, out-degree of node B is 2, so $P\left(1^{2,out}\right)=0.75$. Then, according to definition 3, we can easily get $c_1=\frac{1}{2}\left(P\left(1^{1,out}\right)+P\left(1^{2,out}\right)\right)=\frac{7}{8}$, $c_2=\min\left\{P\left(1^{1,out}\right), P\left(1^{2,out}\right)\right\}=\frac{3}{4}$. Obviously, the greater consistency index $P\left(1^{\alpha,out}\right)$ value, the clearer of travelers' direction in each station is. Therefore, the consistency index $P\left(1^{\alpha,out}\right)$ can describe the travelers' direction under temporal window and more intuitively reflect the traveler's instantaneous traveling purpose

Remark 1. The consistency factors c_1 and c_2 depict the overall level and volatility of the consistency index $P\left(1^{1,out}\right)$ in time window network $G_B^\alpha\left(V,E\right)$ of the temporal network $\left\{G_B^\alpha\left(V,E\right)\right\}_{\alpha=1}^{N}$ (See, section 3.2.2.). In general, the consistency index $P\left(1^{\alpha,out}\right)$ can be further extended to $P\left(k^{\alpha,out}\right)$, and the corresponding definition 2 can also be extended to more general cases.

Definition 4. 1) The time interval $[a,b]$ is divided according to the principle of equal flow (i.e. the number of borrowed records in each time window is equal), and we obtain time windows $\{[t_\alpha, t_\alpha + \Delta t_\alpha]\}_{\alpha=1}^{N}$ that satisfies the definition 3 and the Assumption 2 which is called the flow-type task window, denote for $F\left(\Delta Q, \{\Delta t_\alpha\}\big|_{\alpha=1}^{N}\right)$(See, Fig. 4 (a)), mark as F-window, a set of non-linear time series $t_0(Q_0), t_1(Q_1), \cdots, t_{N-1}(Q_{N-1}), t_N(Q_N)$ is obtained，called as a time series of equal flow.

2) The time interval [a, b] is divided according to the principle of equal time interval, and we obtain time windows $\{[t_\alpha, t_\alpha + \Delta t]\}_{\alpha=1}^{N}$ that satisfies the definition 3 and the Assumption 2 which is called the time-type task window, denote for $F\left(\Delta t, \{\Delta Q_\alpha\}\big|_{\alpha=1}^{N}\right)$ (See, Fig. 4 (b)), mark as T-window. At this time, a set of nonlinear flow sequence $Q_0(t_0), Q_1(t_1), \cdots, Q_{N-1}(t_{N-1}), Q_N(t_N)$ is obtained, called as flow sequence of equal time.

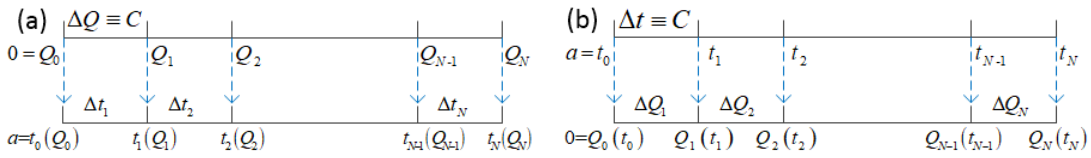


Fig. 4 Task window (a) the flow-type task window; (b) the time-type task window

Remark 2. F-window and T-window are two different types of rebalancing task window. At present, researches on the division of task window is mainly applied to T-window (Regue and Recker, 2014, Shui and Szeto, 2018). The task window of F-window is a new attempt in this paper. As to real-time rebalancing, F-window takes flow linear growth into account, while T-window takes time linear growth. F-window is more suitable as a task window than T-window, and their further comparative study will be given in 3.2.1.2.

The characteristics of instantaneous demand changes under F-window are obvious. During the morning and evening peak period, the number of borrowed (returned) records is more than 500 per minute, and the instantaneous demand of the rebalancing fluctuates greatly (see Figure 10). So the traditional rebalancing under T-window cannot match the speed of such demand changes. At night, static rebalancing is needed to match the supply at the end of day and demand of next day. F-window shows that the time intervals for the same of borrowed (returned) number are not equal (see Figure 12). Each window is usually 2 to 8 minutes in length, with a minimum of 1 minute at the peak and a maximum of 3 hours at night. Although the change in the rebalancing scale due to the same flow is small, the interval of the task window varies greatly.

F-window is a new time window division method in temporal network. F-window can reflect the temporal behavior of travelers better than T-window in details, and is more suitable for the rebalancing in time. It can be seen that B-BN and R-BN have some differences in F-window division. For the convenience, this paper will divide flow-type task windows in the B-BN.

Definition 5. Let $Q_i^{\alpha,out}$, $Q_i^{\alpha,in}$ respectively represent the number of bikes borrowed and

returned from station V_i during the task window $[t_\alpha, t_\alpha + \Delta t_\alpha]$. The initial allocation of storage point is called $Q_{i,0}$ which is determined by two parts: the number of docks E_i and storage point configuration in each station:

$$Q_{i,0} = \frac{1}{2} E_i + \max\left\{\Delta Q_i^\alpha\right\}, \text{ where } \Delta Q_i^\alpha = \sum_{k=1}^{\alpha}\left(Q_i^{k,out} - Q_i^{k,in}\right) \tag{2}$$

In Nanjing of China, the storage points of the stations are built on the surface near the bicycle station. Due to the limited space of this paper, research on the use of critical storage points instead of widely distributed storage points will be presented in other paper.

2.1.3 Management areas based on community detection

Definition 6. In order to finish rebalancing in limited task windows, and reduce the calculative complexity, we divide the whole system into different clusters by Fast Newman algorithm and geographic information. The result of different rebalancing clusters is called the management areas. The management area division is beneficial to the implementation of rebalancing problem.

The Fast Newman algorithm was proposed to divide the community in networks and suitable for large-scale network (Newman, 2004). And its key concept is modularity. Fortunato (2010) has also given the expression of modularity Q under directed weighted network:

$$Q = \frac{1}{W}\left(w_{ij} - \frac{s_i^{out} s_j^{in}}{W}\right)\delta\left(C_i, C_j\right) \tag{3}$$

Where C_i, C_j denote communities of node (station) V_i and node (station) V_j respectively. If the two nodes $\left(V_i, V_j\right)$ belong to the same community, the $\delta\left(C_i, C_j\right)$ value is 1, otherwise is 0.

Here $W = \sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}$ represents the sum of out-strength of all nodes in the network, and $s_i^{in} = \sum_j a_{ji} w_{ji}$, $s_i^{out} = \sum_j a_{ij} w_{ij}$ represents the out-strength and in-strength of the station V_i, w_{ij} represents the weights on the edge E_{ij}, E_{ij} represents the edges from station V_i to V_j, a_{ij} is an adjacency matrix element. If station V_i has connection with station V_j, then $a_{ji} = 1$, otherwise $a_{ji} = 0$ (Zhang et al., 2016).

2.1.4 Starting threshold of the station rebalancing

Definition 7. In the dynamic rebalancing, early value of station inventory is called the station rebalancing starting threshold. In order to enable travelers to borrow and return bikes continuously, operators need to keep the inventory of each station in an interval $[C_{\min}, C_{\max}]$, which is known as the safe stock interval, $C_{\min}, C_{\max}$ are represent the upper and lower limits of the safe stock rate

respectively.

When the number of bikes or empty docks at the bike-sharing system station exceeds the threshold $[C_{\min}, C_{\max}]$, rebalancing is performed at the station under Assumption 1.

2.1.5 Assumptions in the rebalancing scheme

Assumption 1. Assumption of spatial instead of time. Suppose each station has an extra storage point abbreviated as a station warehouse.

During the peak time, the length of task window for rebalancing is very short (generally 2-8 minutes, as shown in fig. 15(a)), so it is difficult to match the whole requirements in time by vehicles. Through setting small warehouse for each station in Assumption 1, it can rebalance directly between the station and the station warehouse to meet the demand in time. By setting up storage points of stations, we can complete the rebalancing scheme directly in each station in time, thus transforming the transportation problem among stations into a problem of setting storage points of the stations, and realizing the idea of 'spatial instead of time'. In subsequent sections, we will study the situation about the key station storage set instead of Assumption 1.

a b

Fig. 5(a) The dock-free bicycle station at the subway exit of Nanjing Olympic Sports Center, where the electronic fence is used as the dock of the bikes on the right side of the station, and the left side is the storage point of the station; (b) A station on Qinglong Street in Nanjing City, where on the right side is a bicycle system with docks, and the left side is the station storage point. Assumption 1 is valid and has practical application background. It has appeared in many cities in China, such as Nanjing and Hangzhou. Station storage points within a station can provide intelligent borrowing and returning services for all vehicles in the station through wireless sensors. This kind of dock-free bicycle station has been operated in Nanjing City of China (See, Fig. 5). This intelligent loan repayment service is used for the station storage point of Assumption 1, which can greatly save the rebalancing time within each station. This is also the original intention of this paper to set the rebalancing time within each station as a unit time. The unit time is set to 1 minute for the convenience in this paper. You can also set it to some other time. As far as we know, almost no one has studied on the rebalancing scheme for bicycles based on this assumption.

Assumption 2. Assumption of rebalancing time within each station. It is assumed that the time of the whole rebalancing process between the station storage point and the dock position is a unit time. In this paper, we assume that the unit time is 1 minute.

During the peak hours, it takes only few minutes to generate a large number of new demands, so the time interval for rebalancing need to be as soon as possible. In the intelligent stations, multiple bikes can be borrowed or returned at the same time. The unlocking or placement of a bike is faster than the loading and unloading from a vehicle. In fact, the vehicle can load and unload a bicycle for

no more than 1 minute (Shui and Szeto, 2018). Therefore, it is reasonable to assume that the rebalancing time within each station is 1 minute.

Definition 8. When performing rebalancing, the ratio of the number of imported bicycles to the station's total docks is *k*, which is called the station's inventory recovery rate.

Assumption 3. Assumption of the inventory recovery rate. According to definition 8, this paper assumes that the inventory recovery rate k is 0.5 (Cyrille et al., 2016； Han et al., 2016)

Remark 3. The relationship between inventory recovery rate k and safe stock interval $[C_{\min}, C_{\max}]$ in station V_i is illustrated in Fig.6 (a). The letter n indicates the number of stations. If k in a station exceeds the range of $[C_{\min}, C_{\max}]$, it means that this station needs to be rebalanced, and the calculation of $[C_{\min}, C_{\max}]$ is described in detail in section 2.2.2.

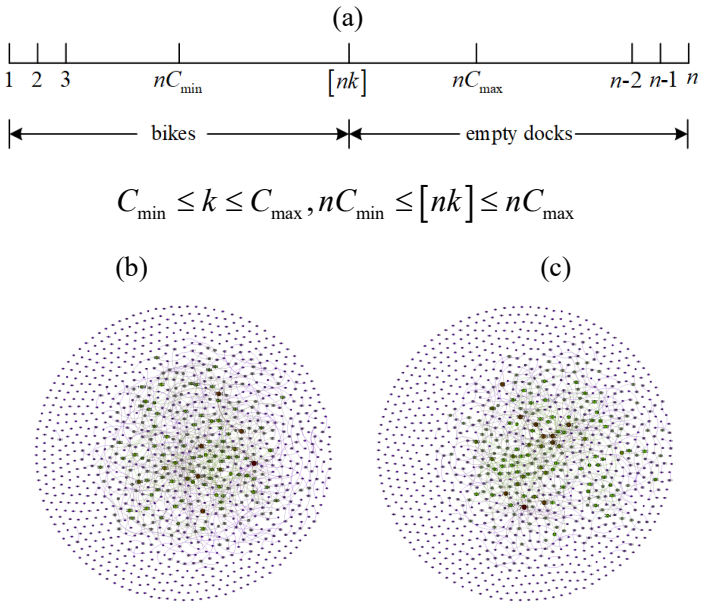


Fig. 6 (a) The relationship between inventory recovery rate and safe stock interval; (b) B-BN topology structure; (c) R-BN between 0:00-6:00 of 2017/3/20

2.2. Methods

Based on the above three assumptions in section 2.1, the traditional rebalancing problem is transformed into a new mixed rebalancing framework, including dynamic rebalancing within each station and static rebalancing among stations. Complex network theory and its topological structure indexes can better reveal the temporal and spatial characteristics of travelers (see Fig.6 (b)-(c)).

2.2.1. Determination of F-window

Determining how to divide the task window is one of the foundations for rebalancing, good task window division can meet the strong time-sensitive demands fluctuation. However there are few literatures related to it. By considering spatial and temporal distribution characteristics of travelers' demand, we first build a task window classification standard by defining the consistency indicators for the direction of travelers, and analyze the appropriate demand increments under each task window. Then we fix the value of the demand increment and calculate all the F-windows (that is rebalancing task window). Then we compare the results with the criteria for window division and Assumption 2. When the result does not meet the window division criterion and Assumption 2, the

flow value is adjusted, and the above steps are repeated, until the optimal flow value is reached, and the optimal task window of rebalancing scheme is reached. Finally, according to the definition 1 and the task window, the initial allocation amount of rebalancing stations is determined. The feedback process is called the temporal distribution learning process (See Fig. 7). The specific steps are as follows:

1. The new standard of window division of temporal network

Consistency factors c_1 and c_2 in definition 3 and rebalancing time within each station in Assumption 2 are used as new criteria for efficient partitioning of B-BN temporal network.

2. Flow increment of window division of temporal network

In order to get the method of dividing the temporal network window that satisfies the definition we use the traffic flow increment ΔQ^{out} of borrowing bikes as a flow increment to divide temporal window of bicycle. And through scenario analysis of the impact of different traffic flow increments ΔQ^{out} on the value of consistency index $P\left(1^{\alpha,out}\right)$ in the timeless network window of borrowing bicycles, the performance of the flow increment ΔQ^{out} is tested.

Remark 4. Previous studies usually used T-window as task window (Regue and Recker, 2014, Shui and Szeto, 2018). This paper introduces F-window for the first time. In section 3.2.1, we also compare these two different division and through analysis we find that F-window with equal flow interval is more suitable for the time-varying of demand and is also more suitable for constructing the dividing standard of task windows.

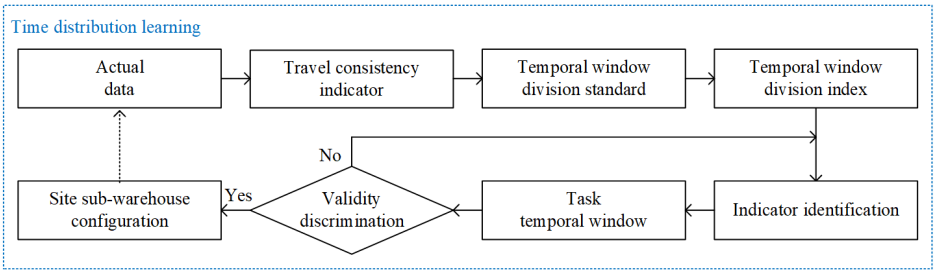


Fig. 7 Data learning process based on temporal distribution

3. Factors determination of temporal window division standard

Combining the results of the evolutionary analysis in the second step, the consistency factors c_1 and c_2 of the Eq. (1) are given.

4. Identification of flow increment

Based on the consistency factors c_1 and c_2, the value of the flow increment ΔQ^{out} of temporal window of bicycle is identified.

5. Task window for rebalancing

According to the numerical value of the flow increment ΔQ^{out} divided by the temporal window,

the temporal network $\left\{G_B^\alpha\left(V,E\right)\right\}_{\alpha=1}^N$ is established and each specific temporal window $\left\{\left[t_\alpha,t_\alpha+\Delta t_\alpha\right]\right\}_{\alpha=1}^N$ is calculated as the task window. When the task window does not meet the Assumption 2, it returns to the third step (see, Fig. 7).

6. The initial configuration of the station storage point

According to Eq. (2), we calculate the initial configuration $Q_{i,0}$ at each station warehouse.

2.2.2. Determination of management area and station starting threshold

The division of management area and the determination of starting threshold for the station are two premises to ensure timely completion of large-scale rebalancing. Therefore, we will build a rebalanced management area by community detection theory, and use the robust analysis method of complex networks to determine the starting threshold for the rebalancing of each station. The validity of community clustering is also known as spatial distribution learning (See Fig. 8).

Community structure is one of the key structural features of complex network, and the interaction between communities determines the overall network performance. Therefore, we give a new method of division of the management area by using the community detection with the directed weighted network and combining the geographical distribution information of stations. The concrete steps are as follows:

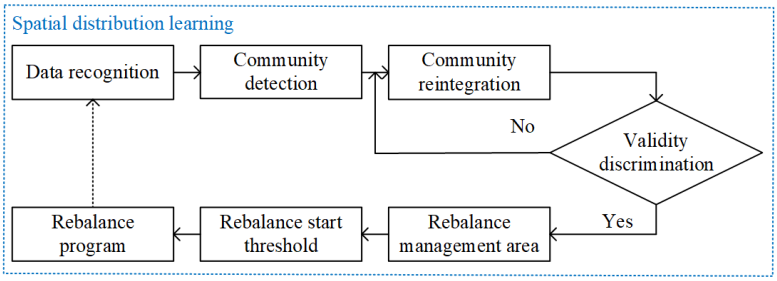


Fig. 8 The data learning process of spatial distribution

1. Selection of network

According to the community detection theory of complex networks, the community structure of different complex networks is different. So this paper selects B-BN for community detection.

2. Determination of the method of community detection

According to the different side operation methods, community detection algorithms can be divided into split method and aggregation method (Newman, 2004). We use Fast Newman algorithm which is suitable for super large-scale network. Among them, the calculation formula of modular degree Q under directed weighted network is given by eq. (3).

3. Reintegration of the community

The bike-sharing system in Nanjing has more than 1000 stations. The stations in different clusters based on community detection can easily form geographical overlap and across. In order to avoid the inconvenience of static rebalancing caused by this problem, we use the geographic information of each station by the ArcGIS software to rationally reintegrate the clusters, compare the different integration results and use the best results as the management area of the bicycle rebalancing.

4. The distance among stations in the management area

Obtaining the actual distance between any two stations is the premise of implementing the static

rebalancing of the bike-sharing system. Therefore, this paper excavates the actual distance information between any two stations in the management area from the Baidu map, and establishes the connection distance matrix between the stations (the latitude and longitude of the bicycle station point is also the Baidu coordinate system.)

5. Station rebalancing starting threshold

The current literature on the safety stock rate of bicycles is rare (Schuijbroek et al., 2017). So we choose robustness theory of complex networks and construct a new calculation method to scientifically analyze the starting threshold of station rebalancing. The process is as follows:

1) Starting threshold of the station $C_{\min}$. Starting threshold of the station $C_{\min}$ is to ensure that vehicles can be borrowed at any station at any time, and can reach the critical value of any station. We use the station as a node and use the number of bikes $q_i(t)$ on the station V_i as the out-degree k^{out}. Each time the system generates a borrowing and returning record, the network corresponds to a directed edge and we construct a directed weighted network $\hat{G}_B(V,E)$. In particular, when $\hat{G}_B(V,E)$ is a fully connected network, the out-degree $k^{out} \geq 1$, it means each station has at least one bike can be borrowed.

2) Starting threshold $C_{\max}$ for station rebalancing. Similar to the $C_{\min}$ analysis process, but we use the number of empty docks $q_i(t)$ on the station V_i (assuming no bicycle could be found at any station) as out-degree k^{out}, so we construct another weighted network $\hat{\mathrm{G}}_R(V,E)$. In particular, when $\hat{G}_R(V,E)$ has no edges, the out-degree $k^{out}=0$, it means no bicycle could be found at any station.

We define that S is the ratio of the number of nodes in the largest connected set to the total nodes in the complex network $\hat{G}_B(V,E)$. According to (Yang et al., 2015), we can obtain:

$$S = \sum P\left(k^{out}\left[1-\left(1-x^{out}\right)^{k^{out}}\right]\right) \tag{4}$$

$$x^{out} = P\cdot\sum\frac{P\left(k^{out}\right)\cdot k^{out}}{\left\langle k^{out}\right\rangle}\left[1-\left(1-x^{out}\right)^{k^{out}-1}\right] \tag{5}$$

Where x^{out} represents the probability of random selection of a node connecting the maximum connected set in the network $\hat{\mathrm{G}}_{\mathrm{B}}(V,E)$. $\frac{P\left(k^{out}\right)\cdot k^{out}}{\left\langle k^{out}\right\rangle}$ represents the probability of random selection of the edge connecting the maximum connected group in the network $\hat{\mathrm{G}}_{\mathrm{B}}(V,E)$, and

$1-\left(1-x^{out}\right)^{k^{out}-1}$ indicates that the probability of having at least one maximally connected group in $k^{out}-1$ directed links (If a traveler rides a bike from V_i to V_j , it means there is a link between V_i and V_j), where k^{out} is the out-degree of the station. According to the Eq. (4) and Eq. (5), we can draw the value of p (the ratio of the number of randomly deleted edges to the total number of edges in the network $\hat{\mathrm{G}}_{\mathrm{B}}\left(V,E\right)$) when $S=1$, that is the starting threshold $C_{\min}$ of station rebalancing we required. Here $1-p$ represents the rate of returning of each station at a certain time. We can draw the value of P corresponding to $S=1$ according to Eq. (4) and Eq. (5) and the value of p is the lower limit of empty dock position for each station, also, the upper limit $C_{\max}$ of the safe stock of the station.

2.2.3. A new rebalancing scheme

Finally we design a new rebalancing framework based on the above temporal and spatial distribution learning process, which is composed of two parts: dynamic rebalancing within each station and static rebalancing among stations (See Fig. 9).

2.2.3.1. A new dynamic rebalancing scheme

The dynamic rebalancing within each station refers to the rebalancing of bicycles from the storage point of each station immediately when the inventory on a station reaches the rebalancing warning value under the assumption 1, and the rebalancing requirements on the station are completed in time. The specific steps are as follows:

First step: establish the F-window task windows based on the temporal network theory and temporal distribution of the travelers demands in the B-BN (see, Definitions 4 and 2.2.1);

Second step: based on the F-window, determine the number of bicycles transferred from storage point within each station (see, Definition 5 and Assumption 1);

Third step: by using the robustness theory of complex networks, figure out the dynamic rebalancing start threshold in the B-BN and the R-BN respectively (see, 2.1.4 and 2.2.2);

Forth step: determine the length of each F-window through the temporal distribution learning process based on Assumption 2;

Fifth step: according to Assumption 3, implement rebalancing in each station and make station's inventory rate recover to 0.5.

2.2.3.2. A new static rebalancing scheme

The purpose of static rebalancing is to relocate bikes at all stations and station storage points to their initial inventory levels to ensure the sustainability of demands. We built a static rebalancing model (See, Table 2) different from previous (Li et al., 2016, Angeloudis et al., 2014) which combines flow distribution model with route planning model and get solution by CPLEX software.

Firstly, this paper assumes that the depot has a certain amount of temporary stock to ensure that the rebalancing in each area can be carried out simultaneously, so as to give priority to the time limit of the task window. According to Table 9, if all vehicles are fully loaded, the inventory needs at most 1,872 bikes which accounts for 1.3% of the average daily usage and averages 1.7 bicycles per station. Secondly, we give the static rebalancing model based on two optimization steps.

The first step is a flow distribution model and generates the rebalancing matrix ($L^{v,\mathrm{M}}$) as an output, aiming at the minimization of the total number of bikes that need to be rebalanced. So we could specify the following problem with the decision variables $l_{ij}^{v,M} \in L^{v,M}$ for each vehicle v in different management area M :

$$Min\sum_{i=0}^{n}\sum_{j=1}^{n+1}\sum_{v=1}^{k} d_{ij}^{v,M} l_{ij}^{v,M} \tag{6}$$

s.t.

$$\sum_{v=1}^{k} y_i^{v,M} = \Delta Q_i^M, \qquad M = 1,\cdots,m \tag{7}$$

$$\sum_{i=0}^{n+1} y_i^{v,M} = 0, \qquad v = 1,\cdots,k \tag{8}$$

$$\sum_{j=0}^{n} l_{ji}^{v,M} + y_i^M = \sum_{j=1}^{n+1} l_{ij}^{v,M}, \qquad i = 1,\cdots,n, v = 1,\cdots,k \tag{9}$$

$$\max\left[0, y_i^{v,M}\right] \le l_{ij}^{v,M} \le \min\left[C + y_i^{v,M}, C\right], \qquad i = 0,1,\cdots,n, j = 1,\cdots,n+1 \tag{10}$$

The objective (6) is to minimize the sum of bikes need to be rebalanced for all management areas, $d_{ij}^{v,M}$ represents the distance between station V_i and station V_j , $d_{ij}^{v,M} = \delta t_{ij}^{v,M}$, here $\delta = 60km/h$, $l_{ij}^{v,M}$ represents the number of bicycles on the vehicle v when it is driven from the station V_i to the station V_j. Constraint (7) indicates that the number of bikes loaded and unloaded by all vehicles in station V_i is equal to the requirements in station V_i , $y_i^{v,M}$ represents the bicycles loaded or unloaded in station V_i by vehicle v, and ΔQ_i^M represents the demand of bicycle in station V_i. Constraint (8) represents that the total amount of loading and unloading bikes by each vehicle is 0, including the number of bikes taken from or brought back to the depot (when $i = 0$ or $i = n+1$, they both mean the depot), constraint (9) requires each station satisfies the flow conservation. Constraint (10) give the upper and lower limits of each vehicle when it is driven among stations, C represents the capacity of each vehicle, and in this paper $C = 48$.

The next step involves the minimization of the travel cost associated with vehicles that rebalance bikes over the whole network. This model is to find the rebalancing paths for carrier vehicles with the decision variable $x_{ij}^{v,M}$. (stations V_i , V_j are linked by vehicle v) when the bicycles controlled by vehicles v is transported from station V_i to station V_j, the value of $x_{ij}^{v,M}$ is 1, otherwise, the value of $x_{ij}^{v,M}$ is 0. In fact, the number of these bikes at each station far exceeds the

capacity of a vehicle. So we argues that each vehicle visits the same station is no longer limited to once.

$$Min \sum_{i=0}^{n}\sum_{j=1}^{n+1}\sum_{v=1}^{k} c_{ij}^{v,M} x_{ij}^{v,M} + \sum_{i=1}^{n+1}\sum_{j=0}^{n}\sum_{v=1}^{k} t^{s} x_{ij}^{v,M} \tag{11}$$

s.t.

$$l_{ij}^{v,M} \le C x_{ij}^{v,M} \qquad i=0,1,\cdots,n,\ j=1,\cdots,n+1,\ v=1,\cdots,k \tag{12}$$

$$\sum_{i=0}^{n} x_{ij}^{v,M} = \sum_{i=1}^{n+1} x_{ji}^{v,M} \qquad j=1,\cdots,n \tag{13}$$

$$\sum_{i=0}^{n}\sum_{j=1}^{n+1} c_{ij}^{v,M} x_{ij}^{v,M} + \sum_{i=1}^{n+1}\sum_{j=0}^{n} x_{ij}^{v,M} t^{s} \le T \qquad v=1,\cdots,k \tag{14}$$

Objective (11) is to minimize the total time cost, including the vehicles' traveling time in all management areas and the total loading and unloading time in each station. t^{s} represents the unit time of loading or unloading, $c_{ij}^{v,M}$ indicates the traveling time cost of each vehicle, $c_{ij}^{v,M} = \lambda t_{ij}^{v,M}$, here $\lambda = 1$. Constraint (12) established the connection between $l_{ij}^{v,M}$ and $x_{ij}^{v,M}$, so that the result of the traffic model can be passed to the route. The constraint (13) represents the conservation of the number of schedules between the stations; and the constraint (14) indicates the rebalancing time limitation by task window T, and this paper selects the first two task windows of each day as rebalancing time (about 0 to 6 o'clock).

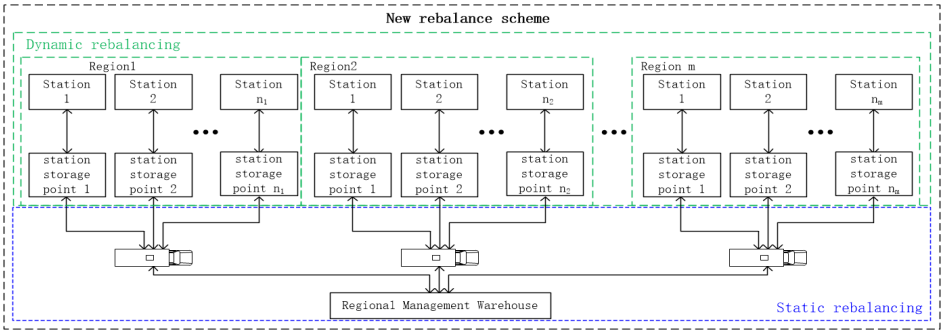


Fig. 9 New rebalancing scheme

3. Empirical study of the new rebalancing framework

3.1. Data preprocessing

The data comes from Nanjing Public Bike Company Limited. There are a total of 980,161 records in main urban areas of Nanjing in China from March 20th, 2017 to March 26th. However, the failure to borrow or return a bike would result in an error log of using time of not exceeding 1 minute. Machine fault and other causes would result in error logs of using time of more than 2 hours (Zhang et al., 2016). In order to avoid the interference of the above two types of error records on practical research, a total of 940,368 valid borrowing records and 940,421 valid returning records have been obtained. The location of stations and actual road information are also obtained.

The whole system's frequency f^{in} , f^{out} and their gap are obtained by borrowing and returning

records per minute (Fig. 10 (a)). We also counted the frequency of borrowing and returning records in each station, and calculated the Relative Entropy between the frequency distribution of borrowing records and the frequency distribution of returning records in each station (Fig. 10 (b)).

Through the above-mentioned frequency evolution, we find that:

1) The frequency of the borrowing and returning records changes drastically with time and shows obvious "tidal" phenomenon (See, Fig. 10): The daily tides are 2 except Wednesday (only 1, affected by rainfall); The two tidal times are near 9:00 am and 6:00 pm respectively, and the tidal peak of the working day is more apparent than the weekend. It means that travelers' demand is strong time-sensitive and inconsistent in time distribution.

2) Although the frequency of borrowing and returning records has a certain similarity in long time distribution, the frequency in each task window is unbalanced because of the strong time sensitiveness. So when we discuss the demand changes under each task window, it is necessary to know that travelers borrow and return the bikes are two completely different behaviors and we should study them separately.

3) According to the evolutionary trend of the Relative Entropy $D_{kld}\left(f_i^{in} \,\|\, f_i^{out}\right)$, it does not tend to be zero. There is a difference in the distribution of records for borrowing and returning. This further illustrates that we should learn them separately.

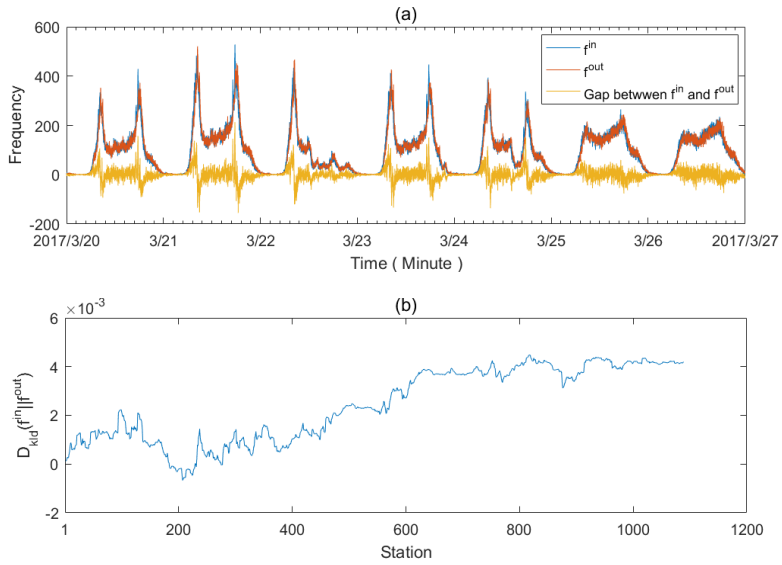


Fig. 10 (a) Temporal distribution of travelers' behavior (The abscissa represents time, and the ordinate represents the number of records within each minute; (b) The Relative Entropy between the frequency distribution of the borrowing and returning bicycles on each station.

3.2. The flow-type task window and the initial value of station storage point

According to the methods established in section 2.2.1., we determine the flow-type task window for dynamic rebalancing and calculate the initial value of the station storage point by data driven.

3.2.1. Evolution analysis of consistency index

3.2.1.1. Consistency index evolution under T-window

The traditional partitioning method for equal time intervals cannot meet the window division criteria newly created of time network in definition 3. In order to verify the above conclusions, we will take different fixed values as the time interval for partitioning temporal window, obtain five sets of temporal networks with equal time intervals, and analyze the change regulation of the

consistency index $P\left(1^{\alpha,out}\right)$ in each group of temporal networks.

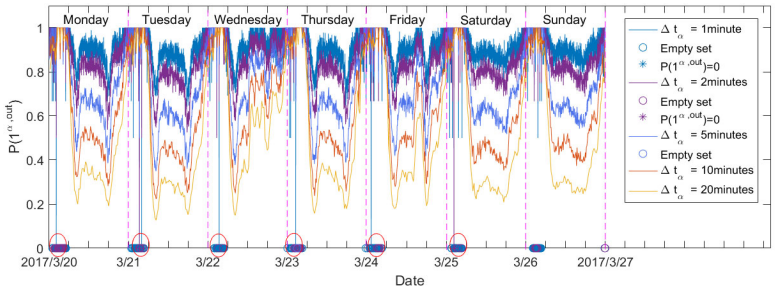


Fig. 11 The evolving map of the consistency index $P\left(1^{\alpha,out}\right)$ under time windows of B-BN

First, we take Δt_α to be 20 minutes and divide one week borrowing records into 504 time windows. Then, according to the records in each time window, the values of the consistency index $P\left(1^{\alpha,out}\right)$ under each time window are calculated. Then we take the time intervals Δt_α as 10, 5, 2 and 1(minute), divide the borrowed records into 1008, 2016, 5040 and 10080 task windows respectively and calculate $P\left(1^{\alpha,out}\right)$ in each group of task windows (See, Fig. 11). We also have sorted out some key features of the consistency index $P\left(1^{\alpha,out}\right)$ (See, Table 3).

Table 3 The value of the consistency index $P\left(1^{\alpha,out}\right)$ in each group of time windows

Δt_α	mean	max	min	range	Number of empty window	Number of $P\left(1^{\alpha,out}\right)=0$
20minutes	0.5318	1	0.1265	0.8735	0	0
10minutes	0.6452	1	0.2236	0.7764	0	0
5 minutes	0.7530	1	0.3492	0.6508	1	0
2 minutes	0.8635	1	0	1	49	2
1 minute	0.9167	1	0	1	331	7

We find that: 1) The $P(1^{\alpha,out})$ of borrowed records always has a large fluctuation (See, Fig. 11), and the values of the range are: 0.8735, 0.7764, 0.6508, 1, 1 respectively (See, Table 3).

2) With the decreasing values of the time window Δt_α, the average value of $P\left(1^{\alpha,out}\right)$ for B-BN constantly increases continuously (See, Table 3), that is $P(1^{\alpha,out}) \propto \frac{1}{\Delta t}$. It means that the smaller the value of Δt_α in B-BN is, the higher the proportion of stations with consistent travel demand within the time window.

3) With the decreasing values of Δt_α, the time windows of the empty set gradually appears:

0→0→1→49→331. At the same time, the number of windows with $P(1^{\alpha,out})=0$ of B-BN are also increasing: 0→0 → 0 → 2 → 7 (See, Table 3), breaking the law of $P(1^{\alpha,out}) \propto \frac{1}{\Delta t}$.

Remark 5. The internal cause of $P(1^{\alpha,out})=0$: as the length of time interval decreases, the amount of borrowing records within a time window is constantly decreasing. For example, when Δt_α is 1 minute, there are only two records in the 130th time window (2:10-2:11, March 20, 2017), from station No. 11204, to the station No. 11148 and No. 11233. Here $P(1^{\alpha,out})=0$. Moreover, There are 2 time windows with $P(1^{\alpha,out})=0$ when Δt_α is 2 minutes and 6 time windows with $P(1^{\alpha,out})=0$ when Δt_α is 1 minute (See, Table 4).

Table 4 The data of $P(1^{\alpha,out})=0$ for the time window

Δt_α	α	$[t_\alpha-\Delta t_\alpha, t_\alpha]$	Q^{Out}	V^{Out}	V^{In}
2minutes	820	3:20-3:22(2017/3/21)	2	15088	12111, 12167
	3668	2:16-2:18 (2017/3/21)	2	11185	11009, 12121
1minute	130	2:10-2:11(2017/3/20)	2	11204	11148, 11233
	1639	3:19-3:20(2017/3/21)	2	15088	12111, 12167
	1679	3:59-4:00(2017/3/21)	2	12180	12121, 12171
	3086	3:26-3:27(2017/3/22)	2	11110	11137, 11207
	4476	2:36-2:37(2017/3/23)	2	12169	12044, 12087
	5845	1:25-1:26(2017/3/24)	2	12168	12044, 12087
	7335	2:15-2:16(2017/3/25)	2	12169	12044, 12087

In summary, the way to obtain Δt_α by dividing equally the time interval is not ideal for the index $P(1^{\alpha,out})$ in time window of B-BN. So we need to find a more suitable method.

3.2.1.2. Consistency index evolution under F-window and comparison with T-window

According to the method of constructing the flow-type task window, we will analyze the results of consistency index $P(1^{\alpha,out})$ under different flow increment ΔQ^{out}.

Firstly, we take the fixed flow increment ΔQ^{out} as 200, and divide total borrowing records into 4702 time windows. Then, the temporal network for borrowing bicycles is established and the values of the consistency index $P\left(1^{\alpha,out}\right)$ under each time window are calculated. We also take ΔQ^{out} as 400, 800, 1200, and 1500 respectively, and records are individually divided into 2351,

1176, 784, and 627 time windows, and consistency index $P\left(1^{\alpha,out}\right)$ of each group of temporal networks is calculated. We obtain the evolution results of the consistency index $P\left(1^{\alpha,out}\right)$ under temporal networks (See Fig. 12). We have sorted out the key values of the consistency index $P\left(1^{\alpha,out}\right)$ in temporal networks under different values of the flow increment ΔQ^{out} (See Table 5).

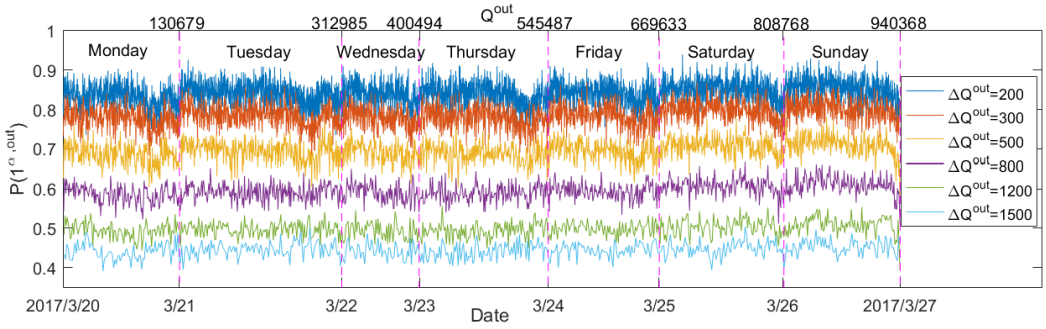


Fig. 12 The evolution map of the consistency index $P\left(1^{\alpha,out}\right)$ under time windows of B-BN

We find that with the increasing values of ΔQ^{out}, the average value of $P\left(1^{\alpha,out}\right)$ for B-BN constantly decreases (See Table 5). Neither the time windows of the empty set nor appears (see Fig. 12). So we have $P(1^{\alpha,out}) \propto \frac{1}{\Delta Q^{out}}$. That means the smaller the value of ΔQ^{out} in B-BN is, the higher the proportion of stations with clear and consistency travel demand in time windows is.

Table 5 The value of consistency index $P\left(1^{\alpha,out}\right)$ with different ΔQ^{out}

ΔQ^{out}	mean	max	min	range
200	0.8414	0.9389	0.7092	0.2297
300	0.7853	0.8694	0.6649	0.2045
500	0.7004	0.7813	0.5849	0.1963
800	0.5953	0.6667	0.5115	0.1552
1200	0.4988	0.5566	0.3937	0.1629
1500	0.4452	0.5038	0.3912	0.1126

So there is a clear correlation between the consistency index $P\left(1^{\alpha,out}\right)$ and the flow increment ΔQ^{out}. In terms of numerical changes, we have $P(1^{\alpha,out}) \propto \frac{1}{\Delta Q^{out}}$. Therefore ΔQ^{out} is a good flow increment to divide task window. Considering the directionality of travel and the number of empty windows, F-window is more suitable for rebalancing than T-window.

Secondly, according to the consistency factor c_1, c_2, under the T-window, when the length of a

task window is getting shorter, the value of the consistency factor c_1 gradually increases from 0.56 to 0.93, and the consistency factor c_2 suddenly drops to zero after experiencing a rapid ascent process (See, Fig. 13(a)); under F-window, when the flow increment of the task window ΔQ^{out} is continuously decreasing. The value of the consistency factor c_1 gradually increases from 0.44 to 0.84, and the value of the consistency factor c_2 gradually increases from 0.39 to 0.71 (See, Fig. 13(b)). Therefore, from the perspective of the consistency factors c_1, c_2, the linear trend of c_1, c_2 under the flow type task window is more obvious than the time type task window. That is, under the flow type task window, the value of the flow increment ΔQ^{out} is easier to determine when the value of consistency factors c_1, c_2 is given. So from the perspective of the consistency factor, F-window is more suitable than the T-window.

Furthermore, from the visibility graph of time series (Lacasa et al.,2008; Tang and Yi, 2017; Zhou et al., 2018), the frequency sequence of borrowing bicycles $\{\Delta Q_\alpha\}|_{\alpha=1}^{N}$ under T-window is a chaotic sequence (Fig.14. (a)-(e)). Here we introduce the Shannon entropy, which is a measure of the degree of system ordering (Shannon, 1948; Ott, 2002).When Δt_α is 1, 2, 5, 10, 20 minutes respectively, the Shannon entropy of the frequency sequence of borrowing bicycles within corresponding task window is 8.7397, 8.0494, 7.1352, 6.4434,5.7526; the sequence $\{\Delta t_\alpha\}|_{\alpha=1}^{N}$ under F-window is also a chaotic sequence (Fig. 14. (f)-(j)). When ΔQ^{out} is 200, 300, 500, 800, 1200 respectively, the Shannon entropy of the duration sequence $\{\Delta t_\alpha\}|_{\alpha=1}^{N}$ of corresponding task window is 7.5485, 7.1497, 6.6660, 5.8451, 5.8451.

It is noted that, under T-window, when the time interval of task windows reaches 1 minute, Δt_α is 1 minute, the Shannon entropy of the frequency sequence of borrowing bicycles $\{\Delta Q_\alpha\}|_{\alpha=1}^{N}$ is 8.7397; under F-window, the minimum time interval of task windows reaches 1 minute, ΔQ^{out} is 500; while ΔQ^{out} is 500, the Shannon entropy of the duration sequence $\{\Delta Q_\alpha\}|_{\alpha=1}^{N}$ of F-window is 6.6660. Therefore, from the perspective of entropy, the chaotic extent of the sequence $\{\Delta t_\alpha\}|_{\alpha=1}^{N}$ under F-window is lower than the frequency sequence of borrowing

bicycles $\{\Delta Q_\alpha\}\big|_{\alpha=1}^{N}$ under T-window. Thus under F-window, the directionality of public travel behavior is more obvious, the corresponding time series is less chaotic, that is, less complicated. Therefore, from the perspective of data complexity, F-window is more suitable than T-window.

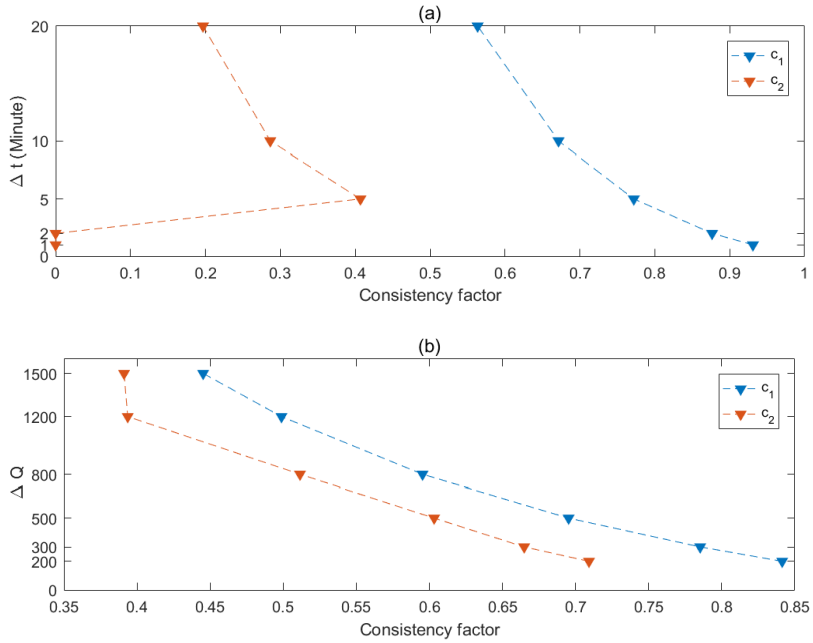


Fig. 13 (a) the evolution of consistency factor under T-window (b) the evolution of consistency factor under F-window

Considering all the above three aspects, we can obviously find that F-window is more suitable than T-window as the task window of the rebalancing scheme.

3.2.2. Flow-type task window

According to the construction method of F-window, combined with analysis in section 3.2.1 and Assumption 2, we take the consistency factors $c_1 \geq 0.7, c_2 \geq 0.6$. According to Eq. (2), we take the flow increment $\Delta Q^{out} = 500$. So the total borrowing records are divided into 1879 task windows for one week. Finally, we establish the temporal network $\{G_B^\alpha(V,E)\}_{\alpha=1}^{N}$, and calculate time interval $[t_\alpha, t_\alpha + \Delta t_\alpha]$ of each task window network $G_B^\alpha(V,E)$ and the corresponding consistency index $P(1^{\alpha,out})$ (See, Fig. 15).

The time interval Δt_α (minute) in the F-window for borrowing records is irregular (See, Fig. 15(a)): There are 8 time intervals more than 150 minutes and the maximum value is 287 minutes; There are 2 time intervals less than 1 minute with the minimum value is 0.9 minutes; There are 1352 time intervals between 2 and 8 minutes, accounting for 71.95% of all task windows; The average value of the interval is 5.3618 (See, Table 7). It means that the criteria for the partition of F-window established by definition 1 is a criterion for irregular task windows $\{\Delta t_\alpha\}\big|_{\alpha=1}^{N}$.

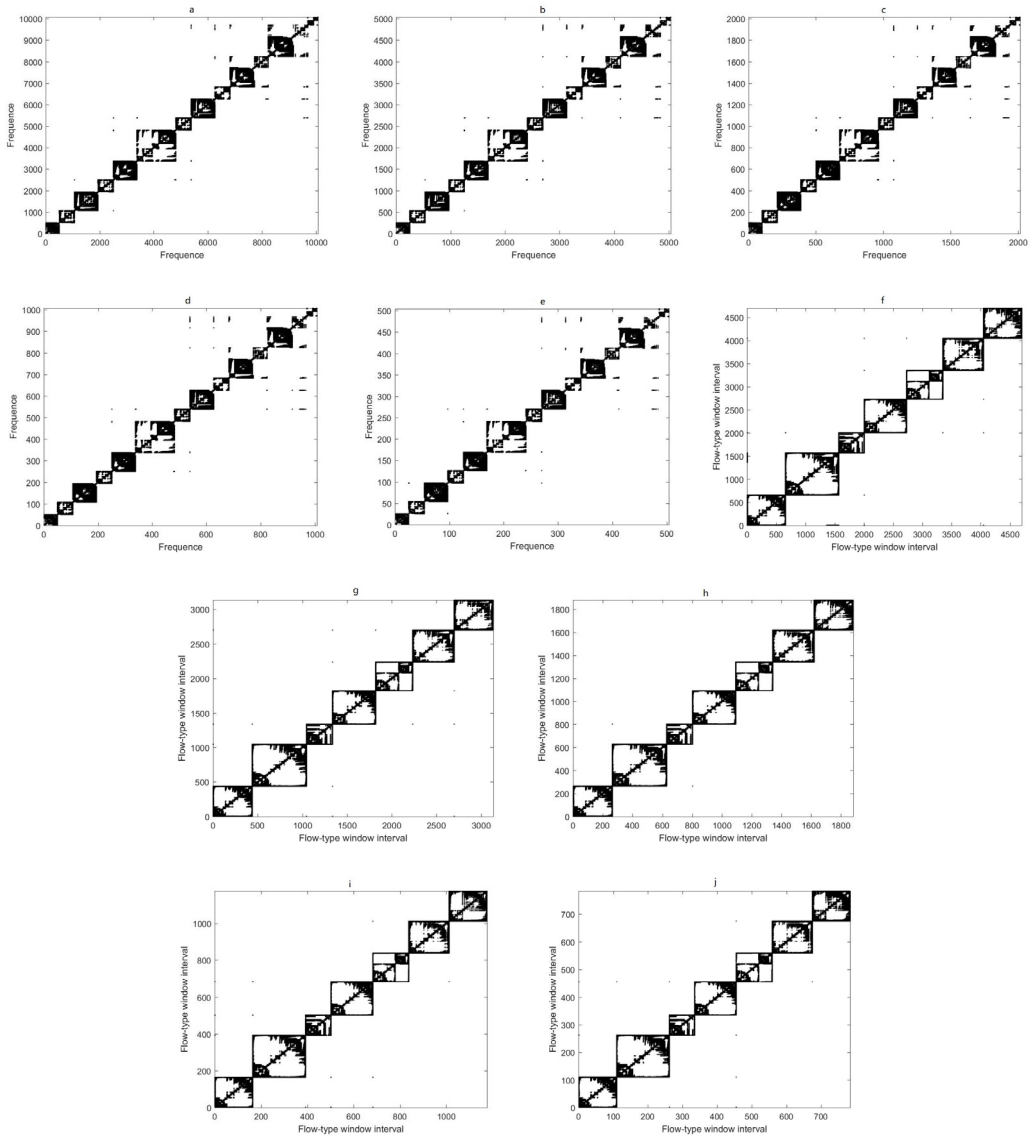
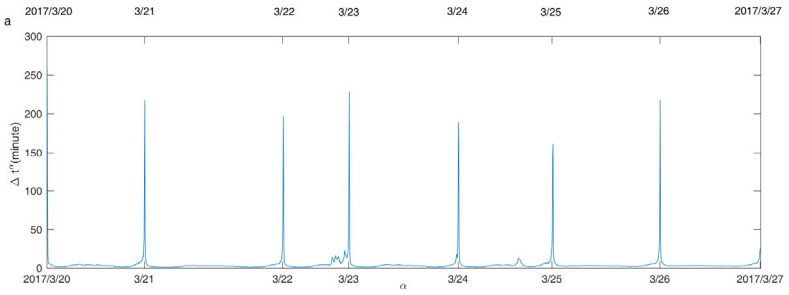


Fig. 14. The frequency sequence visibility graph for borrowing bicycles under two types of task window (a) $\{\Delta Q_\alpha\}\big|_{\alpha=1}^{N}$ with $\Delta t_\alpha=1$; (b) $\{\Delta Q_\alpha\}\big|_{\alpha=1}^{N}$ with $\Delta t_\alpha=2$;(c) $\{\Delta Q_\alpha\}\big|_{\alpha=1}^{N}$ with $\Delta t_\alpha=5$;(d) $\{\Delta Q_\alpha\}\big|_{\alpha=1}^{N}$ with $\Delta t_\alpha=10$;(e) $\{\Delta Q_\alpha\}\big|_{\alpha=1}^{N}$ with $\Delta t_\alpha=20$; (f) $\{\Delta t_\alpha\}\big|_{\alpha=1}^{N}$ with $\Delta Q^{out}=200$;(g) $\{\Delta t_\alpha\}\big|_{\alpha=1}^{N}$ with $\Delta Q^{out}=300$;(h) $\{\Delta t_\alpha\}\big|_{\alpha=1}^{N}$ with $\Delta Q^{out}=500$;(i) $\{\Delta t_\alpha\}\big|_{\alpha=1}^{N}$ with $\Delta Q^{out}=800$;(j) $\{\Delta t_\alpha\}\big|_{\alpha=1}^{N}$ with $\Delta Q^{out}=1200$.



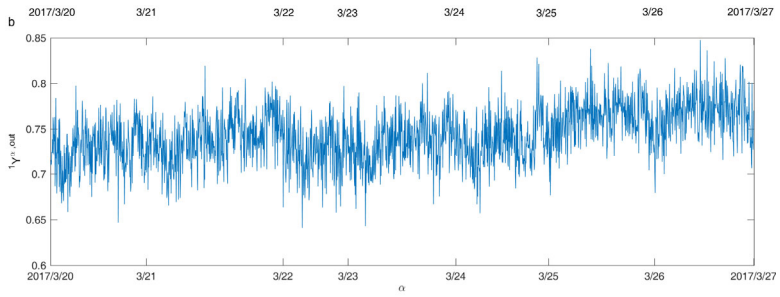


Fig. 15 (a) The time interval of each task window; (b) The consistency index of each task window

There are 8 time intervals $[t_\alpha, t_\alpha + \Delta t_\alpha]$ with rare records early in the morning, and their time intervals exceed 150 minutes (See, Fig. 15(a), Table 7). The number of these windows are 1, 262, 263, 627, 801, 1091, 1340, 1618, and the corresponding time intervals are 0:00:00-4:47:32 (March 20), 0:00:00-2:58:56 (March 21), 2:58:56-5:42:39 (March 21), 1:57:22-5:40:59 (March 22), 0:00:00-4:18:25 (March 23), 0:00:00-3:58:54 (March 24), 1:34:16-5:17:54 (March 25), 1 :37:59-5:23:17 (March 26) respectively. The two time intervals $[t_\alpha, t_\alpha + \Delta t_\alpha]$ with Δt_α less than 1 minute have occurred in the evening peak period (See, Fig15 and Table 6). They are 534th and 535th windows (17:41:12-17:42:08, March 23 and 17:42:08-17:43:07, March 23).

Although the time interval $[t_\alpha, t_\alpha + \Delta t_\alpha]$ of temporal network has a large fluctuation, the value of the consistency index $P\left(1^{\alpha,out}\right)$ within each F-window is relatively stable (See, Fig. 15(b)). Therefore $\Delta Q^{out} = 500$ is a good flow increment to divide task windows in temporal network.

To sum up, the overall dynamic time-varying law of demand can be well depicted through task windows of temporal network and obtained through flow increment. This shows that the task windows division established in the definition 4 is a new standard with irregular time intervals.

Remark 5. The acquisition of the task window is obtained through time-distribution learning. When the fixed flow value is greater than the optimal value 500, the task window is difficult to meet the consistency factor $c_1 = 0.7, c_2 = 0.6$; when the fixed flow value is less than the optimal value 500, the task window is difficult to satisfy the condition of the Assumption 2. The data in the task window with fixed flow values of 1500 and 300 is shown in the Table below (See, Table 6).

Remark 6. The F-window can decompose the total rebalancing demand of the bike-sharing system into each task window relatively consistency, effectively cracking the dramatic fluctuating rebalancing demand brought by the 'tidal' phenomenon of the bike-sharing system, and is conducive to the implementation and completion of the dynamic rebalancing scheme.

Table 6 Task window parameters under different flow increment ΔQ^{out}

ΔQ^{out}	$P\left(1^{\alpha,out}\right)$		Δt_α
	Mean	Min	Min
1500	0.4452	0.3912	3.0333

500	0.7004	0.6033	1.0003
300	0.7853	0.6649	0.5500

Combining with the rebalancing definition in Section 2, we merge the first two task windows each day to form a static rebalancing task window, and finally we get 1872 rebalance task windows in a week. (See, Table 7).

3.2.3. The initial value of the station storage point

The practically utilization number of bicycles is a key factor determining the initial value of the station storage point. It is noted that the month of the most bicycle use in Nanjing during the year is March. The resident population of Nanjing is mainly the working class, and the working cycle is mostly one week. So we choose borrowing and returning records in a week between March 20th, 2017 and March 26th, 2017 in Nanjing to determine the initial value of each station storage point.

Table 7 the rebalancing task windows for between March 20th, 2017 and March 26th, 2017

Number	Starting time	Finishing time	Duration	
1	2017/3/20 00:00:00	2017/3/20 06:07:49	367:49'	Static
2	2017/3/20 06:07:49	2017/3/20 06:24:45	16:56’	
3	2017/3/20 06:24:45	2017/3/20 06:34:01	9:16’	March 20
…	…	…		Dynamic
260	2017/3/20 22:58:02	2017/3/20 24:00:00	61:58’	
261	2017/3/21 00:00:00	2017/3/21 05:42:39	342:39’	Static
262	2017/3/21 05:42:39	2017/3/21 06:09:27	26:48’	March 21
263	2017/3/21 06:09:27	2017/3/21 06:22:29	13:02’	
…	…	…		Dynamic
623	2017/3/21 23:02:07	2017/3/21 24:00:00	57:53’	
…	…	…		
1611	2017/3/26 00:00:00	2017/3/26 05:23:17	323:17’	Static
1612	2017/3/26 05:23:17	2017/3/26 06:06:29	43:12’	March 26
1613	2017/3/26 06:06:29	2017/3/26 6:24:26	17:57’	
…	…	…		Dynamic
1872	2017/3/26 23:00:06	2017/3/26 24:00:03	59:57’	

According to Section 3.2.2, the borrowing amount ${}^{d}Q_i^{\alpha,out}$ and the returning amount ${}^{d}Q_i^{\alpha,in}$ at station V_i within each task window are counted. The cumulative fluctuation of the flow difference at each station $\Delta {}^{d}Q_i^{\alpha}$ is calculated by Eq. (2). We obtain the maximum, the minimum and mean of the fluctuation $\Delta {}^{d}Q_i^{\alpha}$ of each station under each task window in a week (See, Fig. 16)

There are 39 stations where the cumulative fluctuation amplitude $\Delta {}^{d}Q_i^{\alpha}$ exceeds 100, accounting for 3.58% of total numbers of stations; There are also 5 stations with an amplitude of more than 200, the number of which are 85, 233, 235, 414, 419, accounting for 0.37% of total number of stations.

We use $\max_{\alpha}\left\{\Delta {}^{d}Q_i^{\alpha}\right\}$ as the daily maximum demand of the station V_i to obtain the initial

value $Q_{i,0}$ of the storage point of the station V_i by Eq. (2) (See, Fig. 17).

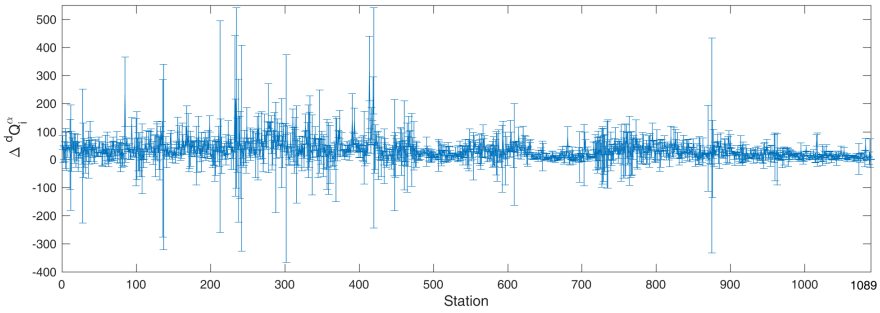


Fig. 16 The amplitude error bar of flow difference of each station in a week (The three values on the error bars are respectively the maximum and minimum of a week's fluctuation amplitude in each task window at stations, and the average of daily station amplitude.)

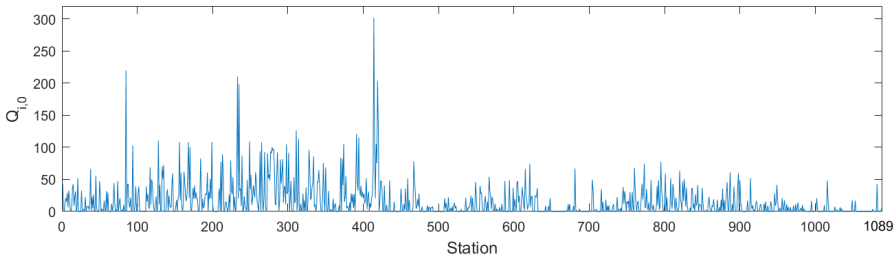


Fig. 17 The initial value of the storage point of the station

It can be seen that operators needs about 58792 bicycles under dynamic rebalancing. 41664 bikes of them are applied to the docks at each station with station inventory recovery rate 0.5, and the remaining are placed in the corresponding station storage points according to the initial value of the warehouse of each station. In fact, there are 468 stations have an initial value of 0, accounting for 42.98% of the total numbers of stations. That means these stations do not require additional arrangements for inventory. However, there are 21 stations with an initial value of exceeding 100, accounting for 1.93% of all and 4 stations with the initial value of exceeding 200, accounting for 0.37% of all. It reflects the importance of these stations and deserves our attention.

Remark 7. The rebalancing scheme proposed in this paper is different to simply expand the number of docks on each station. The above analysis shows that 42.98% of the stations do not need to place bicycles in advance, the storage point needs less space than the dock position.

3.3. The spatial-distribution learning of bike-sharing system

In this section, we study the spatial distribution of bike-sharing system: we mainly use the community detection method and administration regions information to divide the bike-sharing system into different clusters in order to rebalance in a limited time, and we call them management areas. Then according to robustness theory of complex networks, the starting threshold for station rebalancing is also determined.

3.3.1. Management areas division

To divide the bike-sharing system into different clusters is a good way to solve the rebalancing problem with a large number of stations. The concepts of community in complex network structures can represent the tightness of connections among stations. Therefore, we will establish rebalancing management areas mainly based on the community detection of complex networks.

3.3.1.1. Community detection

We first select B-BN for the community division. Based on the Fast Newman community detection algorithm (Newman, 2004) and combined the modularity Eq. (3) of the directed weighted network, we have obtained a community division of B-BN. The division makes 1089 bicycle stations into 15 communities (clusters). The number of stations in each community is: 155 (bright red), 152 (orange), 125 (bright yellow), 106 (apple green), 91 (chalcedony green), 87 (blue), 82 (fir green), 76 (purple), 46 (maroon), 45 (deep purple), 45 (gray), 35 (pink), 30 (light green), 8 (brown) and 6 (dark blue) (see, Fig. 18).

3.3.1.2. Community reintegration

Although the above-mentioned communities have a modularity of 0.6509 and can well characterize the overall structural characteristics of B-BN. However, there are obvious overlaps and intersections among the stations of the communities, not conducive to the implementation of rebalancing. Therefore, we need to rationally re-integrate the aforementioned communities based on the actual geographical location of the bicycle stations.

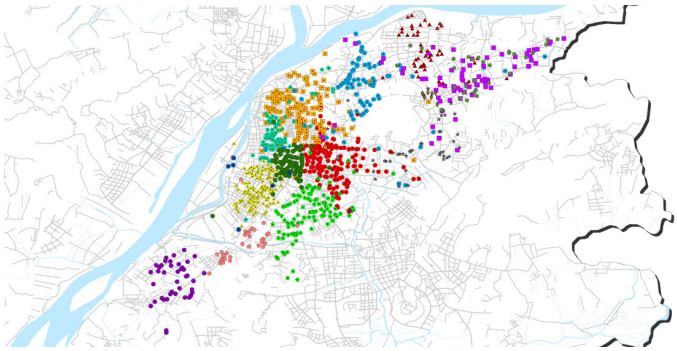


Fig. 18 Community detection results of B-BN

Based on the geographical distribution of bicycle stations in Nanjing and combined with the rebalancing experience (under the principle of least overlapping between communities) of Nanjing Bike-Sharing Company, we reintegrate the above 15 communities and obtain five regions of rebalancing (See, Fig. 19). The number of bicycle stations in each region is: 261 (red), 325 (green), 131 (yellow), 80 (blue), 292 (purple).

Remark 8. The determination of the management area based on community detection is a process of continuous learning, repeated comparison, and selection of the optimal community reintegration. The figure below is an integration scheme that we abandoned in the spatial distribution learning process (see Fig. 20). In the shadowed part of the scheme, there is a clear overlap of communities, which is not conducive to the implementation of the rebalancing scheme of the bike-sharing system.

3.3.2. Characteristics of distance among stations

The distance information between any two stations is the premise and basis for implementing static rebalancing. For this purpose, we use the web crawler to obtain the actual distance between the two stations in each management region from the Baidu map software, and calculate the maximum, mean, and minimum distance from each station to other stations in the management region (See Fig. 21).

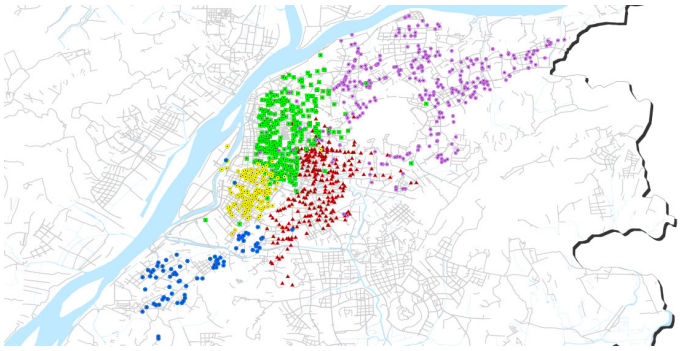


Fig. 19 The rebalancing management areas of bike-sharing system

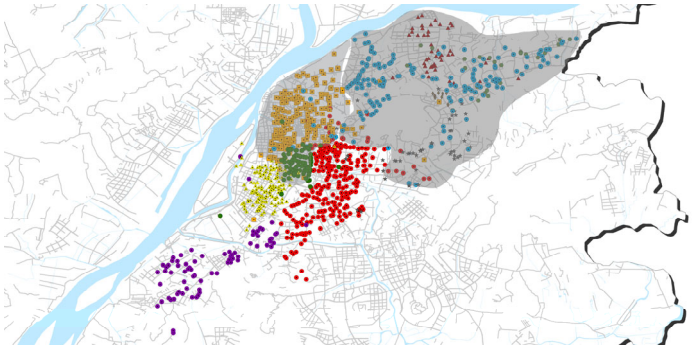


Fig. 20 The rebalancing management areas of bike-sharing system

From the above analysis, it can be found that the fifth management region has the maximum mean distance between two stations in the regions, with the value being 15.1806; The second management region ranks second with the value of 11.4012; The third and fourth positions are taken by the fourth and first management regions, the values of which are 7.4035 and 6.8119 respectively; The third management region has the minimum average distance between two stations in the management regions, and its value is 3.7694 (See, Fig.21).

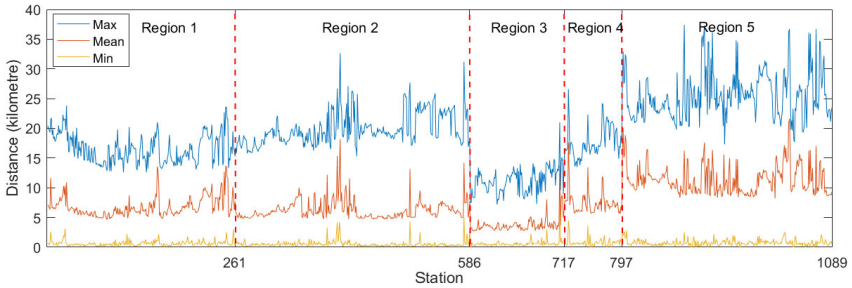


Fig. 21 Distance among stations in the management region

According to the urban traffic regulations in Nanjing, the speed limit is 60 km/h. So even if the loading and unloading time of the bike and the waiting time of the traffic light are not considered, the travel time between two stations in each management region respectively reaches 11.3855, 8.5509, 5.5526, 5.1089 and 2.8270 minutes. However, the analysis results in Section 3.2 tell us that the time intervals of 1371 windows in the 1881 task windows range from 2 to 8 minutes, accounting for 72.89% of the total number of task windows. Therefore, both the static rebalancing and dynamic rebalancing are hard to meet the rebalancing requirements of the bike-sharing system in time, and the phenomenon that no bike can be borrowed and no dock for the returning bikes cannot be solved by increasing the dispatching vehicle and optimizing the dispatching route. In fact, the method of

spatial instead of time method proposed in this paper is an effective solution to this problem.

3.3.3. The starting threshold for station rebalancing

Rebalancing cannot be completed in time is one of direct factors that supply cannot match demand. Therefore, in this section, we will use the robustness theory of the complex network to give the starting threshold of the bike-sharing system rebalancing scheme and determine when a station needs to be rebalanced.

According to section 2.2.2, the starting threshold of station can be obtained from the upper bound $C_{\max}$ and lower bound $C_{\min}$ of a station's safety stock. In order to determine lower bound $C_{\min}$, we use the station as a node and the dock numbers of a station as out-degree of this node to establish a directed-weighted network. Then we randomly select a station V_i , gradually delete some of edges E_{ij} on the node V_i and the deletion ratios are respectively: 1, 0.95, 0.90, 0.85, 0.80,$\cdots$. The corresponding values of S are obtained according to Eq. (4) and (5). We observe the relationship between the values of S over the values of p (See, Fig. 22).

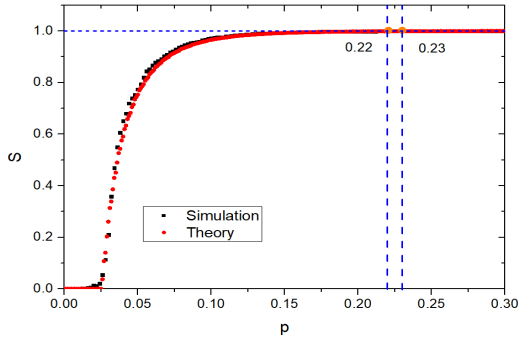


Fig. 22 The starting threshold of the station rebalancing

From Fig. 22, we can see that $P = 0.225$ is the theoretical value of the minimum limit $C_{\min}$ of safety stock at a station. It means that the sum of bicycles $q_i(t)$ available on station V_i and the sum of empty docks $E_i - q_i(t)$ for returning is constant, equal to the maximum number E_i of bicycles containable in the station V_i. So we can get $C_{\max} = 1 - C_{\min}$. That is, 0.775 is the theoretical value of the maximum limit $C_{\max}$ of the station safety inventory. Based on the operating experience of operators, we confirm that $C_{\min} = 0.2$, $C_{\max} = 0.8$ are the minimum and maximum limits of station's safety stock . The value 0.225 and 0.775 obtained in this paper are more accurate than the value 0.2 and 0.8.

3.4. Empirical Studies

In this section, we will conduct an empirical analysis of the learning rebalancing scheme of the bike-sharing system based on the data from March 20th, 2017 to March 26th, 2017 in Nanjing.

3.4.1. Dynamic rebalancing within each station

According to the implementation method of the rebalancing scheme (See, section 2.2.3.1), we first calculate the task window for each rebalancing among stations (See, Table 5), and count the number of available bicycles and empty docks at all stations within each task window. Combining with the starting threshold of each station in section 3.3.2, we determine the rebalancing requirements of each station in each task window as well as the cumulative rebalancing requirements of each management region over time (See Fig. 23). Then the dynamic rebalancing frequency of each station is also counted (See Fig. 24). Finally, the dynamic inventory of each station is sorted out (see Fig. 26).

Based on the actual data, we find:

1) Although the number of the bicycle stations in Nanjing is very large, reaching 1089. But the number of stations that need to be rebalanced in diffferent management region under each task window is small. The maximum value is only 19 (See, Fig. 23(a)). Therefore, the division of task windows and management regions can effectively relieve peak pressure.

2) Judging from the ranking of dynamic rebalancing frequency (See, Fig. 23(b)): In the region 2, the scale of stations to be rebalanced under the task window is the largest, and the total number of stations that need to be rebalanced within a week is also the largest. Although the size of the dynamic rebalancing stations between region 1 and region 3 under each task window is inconsistent, the total number of stations requiring dynamic rebalancing within a week is basically the same. The last two are the region 5 and region 4 respectively.

3) According to the frequency of dynamic rebalancing within each station (See, Fig. 24), there are 63 stations with value of 0, accounting for 5.79% of all stations; There are 642 stations with a value of less than 25, accounting for 58.95%; There are 357 stations with a value of between 25 and 100, accounting for 32.78% of all stations; Only 27 stations value of exceeding 100, accounting for 2.48%; The largest frequency occurs on the station No. 420 and reaches 255.

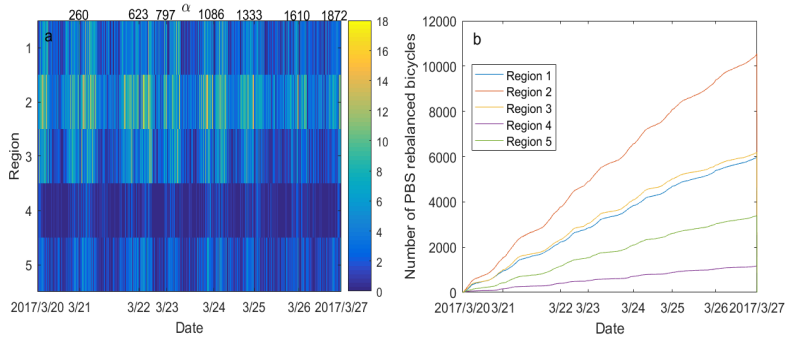


Fig. 23 (a) The number of stations that need to be rebalanced in each area under the dynamic rebalancing windows; (b) Rank of dynamic rebalancing frequency in different regions

4) The dynamic rebalancing at station No. 420 (See, Fig. 25) mainly occurred in the morning and evening peak periods, including 140 times to the storage point and the maximum number of bikes is 24 and the average is 15.8. There are also 115 rebalancing from the storage point to station No. 420, the maximum number of bikes is 29 and the average is 16.8.

5) Through the dynamic rebalancing within each station, we ensure the stations' average inventory fluctuates around 0.5, and maximum and minimum values are always in the interval $\left[C_{\min}, C_{\max}\right]$ (See, Fig. 26).

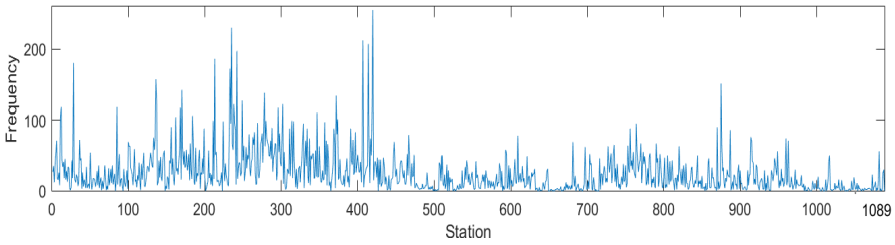


Fig. 24 The dynamic rebalancing frequency of each station

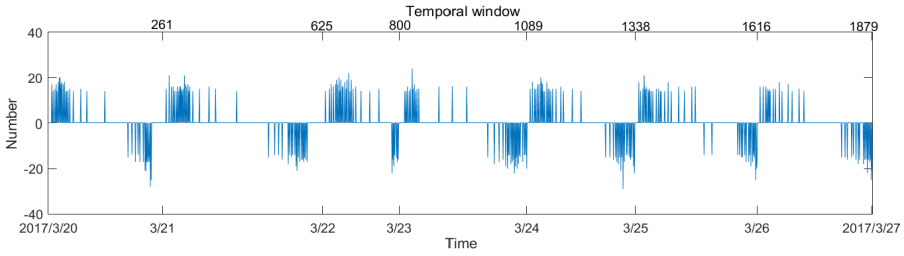


Fig. 25 The amount of rebalancing fluctuation in station No. 420

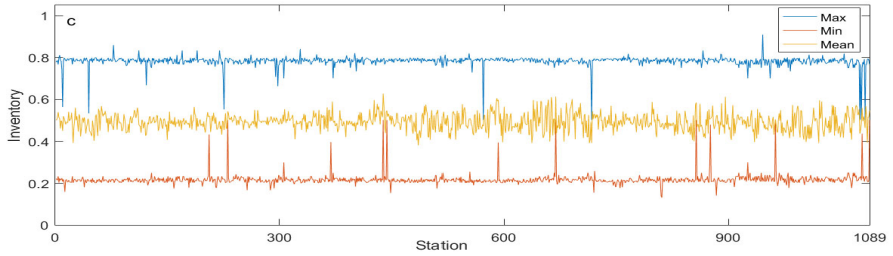


Fig. 26 The dynamic inventory of each station in the dynamic rebalancing window

The above analysis shows that based on Assumption 1 (spatial instead of time), dynamic rebalancing within each station can match the strong time-sensitive changes of demand in a short time. In fact, the time of previous dynamic rebalancing methods generally took more than one hour even if vehicles will be increased or the vehicle route is continuously optimized (Contrado et al., 2012, Zhang et al., 2017), it is too difficult to develop a dynamic rebalancing scheme in 1 minute time scare because the average traveling time between stations is between 2-8 minutes (See, Fig. 21). Therefore, the method of spatial instead of time proposed in this paper can effectively shorten the time of rebalancing and effectively match the supply to the demand .

3.4.2. Static rebalancing among stations

Static rebalancing is the basis for ensuring dynamic rebalancing within each station. When the system scale is large, the time of static rebalancing can last for several hours (Forma, et al., 2015). It is also closely related to the number of vehicles and the capacity of each vehicle (Ho and Szeto, 2017). So the model in this paper can make the large-scale system complete static rebalancing within the task window of 6 hours before dawn. In order to ensure the static rebalancing of the five regions at the same time, we assume that a total general warehouse is set up in Nanjing, and at least 1,872 bicycles are placed in the depot, accounting for 1.3% of the average daily usage, with average 1.7 bicycles per station. In fact, Nanjing has such a general warehouse.

First, we divided the station into 5 management areas by Definition 6 (See，section 3.3.1.). Then we obtain the task window for daily static rebalancing (See, Table 8). The average length of windows is 342 minutes. The maximum length is 367 minutes, appearing in 0:0:0-6:07:49 of March 20th, 2017 and the minimum length is 317 minutes, appearing in 0:0:0-5:17:43 of March 25th, 2017.

Table 8 The task window for daily static rebalancing

Date	Task window	Window duration(Minute)
2017.3.20	0:0:0-6:07:49	367.82
2017.3.21	0:0:0-5:42:39	342.65
2017.3.22	0:0:0-5:40:41	340.68
2017.3.23	0:0:0-5:53:05	353.08
2017.3.24	0:0:0-5:51:21	351.35
2017.3.25	0:0:0-5:17:43	317.71
2017.3.26	0:0:0-5:23:13	323.21

Table 9 The number of vehicles for rebalancing

Date	Region 1	Region 2	Region 3	Region 4	Region 5	total
2017.3.20	8	8	3	3	8	30
2017.3.21	9	9	3	3	7	31
2017.3.22	8	11	6	4	10	39
2017.3.23	8	7	3	3	8	29
2017.3.24	5	7	3	2	8	25
2017.3.25	8	9	3	2	8	30
2017.3.26	6	6	4	2	8	26

Table 10 The total travel distance of vehicles

Date	Region 1	Region 2	Region 3	Region 4	Region 5	total
2017.3.20	2104.2	2114	722.7	639.2	2210.4	7790.5
2017.3.21	2334.2	2301.3	625.8	834.92	1836.6	7932.8
2017.3.22	1851.1	2718.6	1456	887.2	2638.7	9552
2017.3.23	2131	1588.3	759.3	710	2206.2	7394.8
2017.3.24	1089.6	1545	583.55	479.54	2046.6	5744.3
2017.3.25	2019.2	2045.5	673.79	488.02	1777.1	7003.6
2017.3.26	1358.8	1268.3	893.7	512.4	1926.1	5959.3

According to the actual traffic conditions in Nanjing, we take the average speed of the vehicle for 60 km/h. At the same time, assuming that the stopping time t^s of the vehicles at each loading station would not change for 1 minute (Zhang et al., 2017). Then the number of bikes that need to be rebalanced for each station in each management region are calculated. Using the static rebalancing scheme (See, Section 2), we obtain the time in each management region for static rebalancing (See Fig. 24), the number of vehicles (See, Table 9 and Fig. 27 (a)) and the total travel distance of the vehicles (See, Table 10 and Fig. 27 (b)).

During the static rebalancing between March 20th, 2017 and March 26th, 2017, we find that:

1) The number of vehicles for rebalancing: the average of the vehicles in each management area is 6, and the maximum is 11 in management area 2 on March 22; the minimum value is 2 in management area 4 on March 24, March 25 and March 26. The average number of vehicles used for one week in each management area is 42 and the maximum value is 57 in management area 2 and management area 5; the minimum of the vehicles used for the week in each management area is 19 in management area 4). The average of the vehicles in all areas is 30; the maximum value is 39 on March 22 and the minimum is 25 on March 24.

2) The total rebalancing distance: the average daily driving distance of vehicles is 1467.9 km, the maximum is 2718.6 km in management area 2 on March 22 and the minimum is 479.5 km in management area 4 on March 24. The average of the total driving distance in each management area is 10275 km, the maximum is 14642 km in the management area 5, and the minimum is 4551 km in management area 4. The average of daily driving distance of management areas is 7339.6, the maximum is 95516 on March 22 and the minimum is 5744.3 on March 24.

3) The time of static rebalancing: Static rebalancing in each management area can be completed in the task windows (See, Fig. 28). The time of static rebalancing in management areas are more than 4 hours, and their average amplitude of fluctuations is 52.07 minutes.

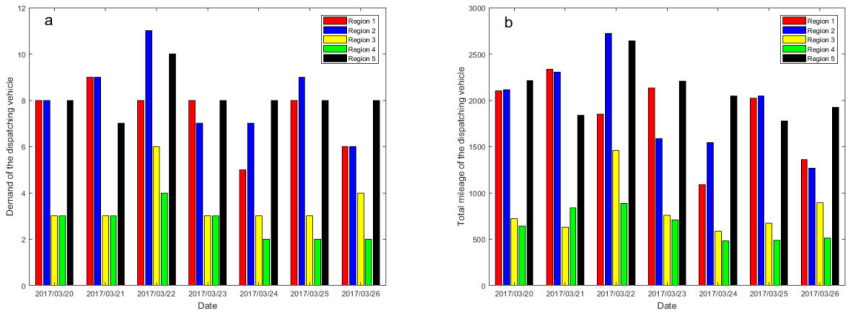


Fig. 27 (a) Daily dispatch vehicle demand; (b) Daily driving total mileage of vehicles

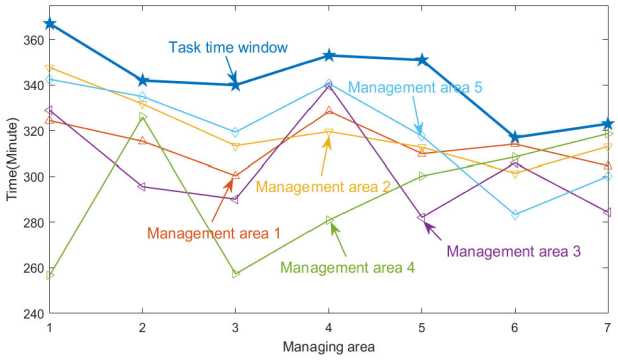


Fig. 28 The time interval of static rebalancing

In summary, based on three assumptions, our rebalancing framework, which combines dynamic rebalancing within each station and static rebalancing among stations, can meet travelers' demand fluctuation in Nanjing well. It effectively solved the problem of large scale stations and uneven distribution of time and space. Through F-window division, the rebalancing scheme in this paper meet the rebalancing requirements of all stations in the bike-sharing system in time, without the phenomenon of no available bike and no dock for the returning bikes.

3.5. Comparison between F-window-based and T-window-based rebalancing scheme

Compared with the traditional dynamic rebalancing, this paper can match the strong time-sensitive changes of demand. The current existing dynamic rebalancing window length is basically fixed (Table 11) and mostly in hours，for example some researchers use 4 hours as a time window for dynamic rebalancing (Kloimullner et al., 2014)，some others also use 3 hours as a time window (Zhang et al., 2017), and the shortest is also 1 hour (Shui and Szeto, 2018; Brinkmann et al., 2019). This would result in a large amount of demand loss during the peak period due to rebalancing cannot be completed in each window. The average length of F-window is only 4.06

minutes (See, Fig. 15). Since the length of task window dynamically matches the demand fluctuation, in the morning and evening peaks, in order to face the sudden surge in demand, the shortest time window is only 1 minute; when the early morning demand is small, the longest window length can reach 59 minutes (See, Table 11). Therefore, the rebalancing within stations matches the strong time-sensitive demand fluctuation well, which is unmatched by traditional dynamic rebalancing.

Table 11 The distribution of the time length of F-windows

Time length of F-windows/minutes	[1,2]	[2,8]	[8,20]	[20,30]	[30,40]	[40,50]	[50,60]
Number of windows	386	1352	106	15	3	3	1

From the perspective of the real-time rebalancing of demand, the solution of this paper far exceeds the traditional dynamic rebalancing regardless of the cumulative station satisfaction frequency and level of demand satisfaction, greatly reducing the demand loss. Traditional dynamic rebalancing ignores real-time changes in demand per hour. Therefore, according to the traditional dynamic rebalancing every 1h, the total number of stations and the total number of bicycles that have been reached balance in one week is relatively small, 37488 and 1228658 respectively. In this paper, based on F-window, the cumulative frequency of stations and the total number of bicycles reached balance are 1,193,083 and 34,787,759 respectively，which is 31.82 and 28.31 times respectively in the case of dynamic rebalancing per hour.

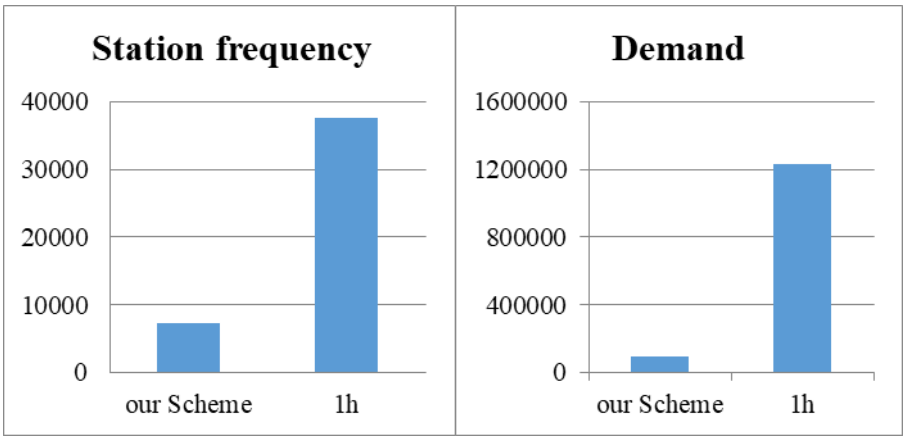


Fig. 29 Stations meet frequency and demand satisfaction under two rebalancing schemes

The static rebalancing in this paper also effectively reduces travel cost. Based on Assumption 1 of spatial instead of time, this paper transforms the system rebalancing requirements in into dynamic rebalancing within each station and completes it in each F-window. Then the static rebalancing of whole system is completed by vehicles in the two task windows before dawn. So the total frequency of visited stations and the rebalanced bicycles by vehicles in a week are 7240 and 96270, respectively account for 0.61% and 0.28% of rebalancing scheme in this paper and reach 19.31% and 7.84% of the hourly dynamic rebalancing scheme (See, Fig. 29). The scheme in this paper requires 17128 more bicycles, accounting for 17.1% of the total existing deployment (website: http://www.nanjing.gov.cn/mszx/201811/t20181102_1077104.html).

4. Conclusion and future research

By learning spatial-temporal distribution of bike-sharing system demand fluctuation in Nanjing, we have built a new rebalancing framework and verified the feasibility of model. Firstly, we give a new flow-type task window (F-window) by defining the consistency index. We find that it is more

suitable for rebalancing than time-type task window (T-window) based on three aspects analysis.

Secondly, through three assumptions, the temporal-distribution learning model including task window and station storage configuration, are built to realize new dynamic rebalancing. The spatial-distribution learning method is introduced to give management areas for static rebalancing, we have divided the whole system into 5 management areas.

Thirdly, through the robustness theory of complex networks, we have defined the safety stock rate of each station and calculated its upper and lower bound in order to clarify the amount of the rebalancing within each station.

Finally, this paper have designed a new framework to solve rebalancing problem which contains two aspects: dynamic rebalancing within each station and static rebalancing among stations.

Compared with the scheme of dynamic rebalancing per hour, we find that the scheme of this paper can match the strong time-sensitive changes of demand than the traditional dynamic rebalancing. The cumulative frequency of station and demand satisfaction in a week is close to 31 times and 28 times the dynamic rebalancing per hour, which greatly reduces the loss of demand. At the same time, the number of stations and the total number of vehicles' visits is only 19.31% and 7.84% of the hourly rebalancing, which effectively alleviates the traveling cost and this paper only need to increase the amount of delivery by only 17% of the existing total.

The research in this paper is based on the analysis of real-time data driven. Due to the limitations of the collected data, the initial allocation of the station storage point, the rebalanced management area and the starting threshold of station rebalancing are all static. It needs further study to determine the cycle of these three variables establish the dynamic mechanism changing over time period and optimize the existing rebalancing strategies. Due to the complexity of the rebalancing scheme, optimization of storage points, research into the use of critical storage points instead of widely distributed storage points will be presented in another paper.

Acknowledgements

The Research was supported by the Major Project of the National Natural Science Foundation of China (71732003, 71690242).

References

Ahmadreza F. I., Robert H., Lavanya M., Naveen E., 2017. An empirical analysis of bike sharing usage and rebalancing: Evidence from Barcelona and Seville. Transport. Res. Part A, 97, 177-191.

Alderson A S, Beckfiled J. Power and position in the world city system. American Journal of Sociology, 2004, 109(4), 811-851

Angeloudis P., Hu J., Bell M. G. H., 2014. A strategic repositioning algorithm for bike-sharing schemes. Transportmetrica A., 10, 759-774.

Barrat A, Barthelemy M, Vespignani A. Modeling the evolution of weighted network[J]. Physical Review E, 2004, 70(6), 1-13.

Banerjee A., Chandrasekhar A. G., Duflo E., 2013.The diffusion of microfinance. Science, 6144, 1236498.

Behzad M. V., Iraj M., Nezam M., Esmaile K., 2020. Balancing Public Bicycle Sharing System Using Inventory Critical Levels in Queuing Network. Computers & Industrial Engineering. https://doi.org/10.1016/j.cie.2020.106277 (in press).

Benchimol M., Benchimol P., Chappert B., Taille A., Laroche F., Meunier F., Robinet L., 2011. Balancing the stations of a self-service “bike hire” system. Rairo-Oper. Res. 45, 37-61.

Brinkmann J., Ulmer M.W., Mattfeld D.C., 2016. Inventory Routing for Bike Sharing Systems. Transport. Res. Procedia 19, 316-327.

Brinkmann J., Ulmer M.W., Mattfeld D.C., 2019. Dynamic Lookahead Policies for Stochastic - Dynamic Inventory Routing in Bike Sharing Systems. Computers & Operations Research, 106, 260-279.

Caggiani, L., Camporeale, R., Ottomanelli, M., Szeto, W.Y., 2018. A modeling framework for the dynamic management of free-floating bike-sharing systems. Transport. Res. Part C, 87, 159-182.

Chemla D., Meunier F., Wolfler C. R., 2013a. Bike sharing systems: solving the static rebalancing problem. Discr. Optim. 10, 120-146.

Chemla, D., Meunier, F., Pradeau, T., Calvo, R.W., Yahiaoui, H., 2013b. Self-Service Bike Sharing Systems: Simulation, Repositioning, Pricing. Hyper Articles en Ligne -HAL [online], hal-00824078. Available at: https://hal.archives-ouvertes.fr/hal-00824078.

Chiariotti F., Pielli C., Zanella A., Zorzi M., 2018. A dynamic approach to rebalancing bike-sharing system, MDPI., 18(2), s18020512.

Contardo C., Morency C., Rousseau L. M., 2012. Balancing a Dynamic Public Bike-sharing System. Technical Report CIRRELT-2012-09, Montréal.

Cyrille M. D. C., Geoffrey C., Isabelle T., 2016. Bike-share rebalancing strategies, patterns, and purpose. J. Trans. Geog., 55, 22-39.

Dell’Amico M., Hadjicostantinou E., Iori M., Novellani S., 2014. The bike sharing rebalancing problem: mathematical formulations and benchmark instances. Omega. 45, 7-19.

Du Y. C., Deng F. W., Liao F. X., 2019. A model framework for discovering the spatio-temporal usage patterns of public free-floating bike-sharing system. Transport. Res. Part C, 103, 39-55.

Erdogan G., Battarra M., Wolfler C. R., 2015. An exact algorithm for the static rebalancing problem arising in bike sharing systems. Eur. J. Oper. Res. 245, 667-679.

Erdoğan G., Laporte G., Calvo. W. R., 2014. The static bike relocation problem with demand intervals. Eur. J. Oper. Res. 238, 451-457.

Feng L., Monterola C. P., Hu Y. Q., 2015. The simplified self-consistent probabilities method for percolation and its application to interdependent networks. New J phys. 17, 1-15.

Forma I. A., Raviv T., Tzur M., 2015. A 3-step math heuristic for the static repositioning problem in bike-sharing systems. Transport. Res. Part B, 71, 230-247.

Fortunato S., 2010. Community detection in graphs. Physics Reports., 486, 75-174.

Han L., Luong B. T., Ukkusuri S., 2015. An algorithm for the one commodity pickup and delivery traveling salesman problem with restricted depot. Netw. Spat. Econ., 16, 1-26.

Ho S. C., Szeto W. Y., 2017. A hybrid large neighborhood search for the static multi-vehicle bike-repositioning problem. Transport. Res. Part B., 95, 340-363.

Kadri A. A., Kacem I., Labadi K., 2016. A branch-and-bound algorithm for solving the static rebalancing problem in bike-sharing systems. Comput. Ind. Eng., 95, 41-52.

Kloimullner C., Papazek P., Hu B., Raidl G. R., 2014. Balancing Bike Sharing Systems: An Approach for the Dynamic Case. European Conference on Evolutionary Computation in combinatorial optimization, 8600, 73-84.

Lin J. H., Chou T. C., 2012. A geo-aware and VRP-based bicycle redistribution system. Int. J. Veh.

Technol., 1-14.
Leonardo Caggiania, Rosalia Camporealea, Michele Ottomanellia, Wai Yuen Szetob, 2018. A modeling framework for the dynamic management of free-floating bike-sharing systems. Transport. Res. Part C, 87, 159-182.
Legros B., 2019. Dynamic repositioning strategy in a bike-sharing system; how to prioritize and how to rebalance a bike station. Eur. J. Oper. Res. 272(2): 740-753.
Li Y. F., Szeto W. Y., Long J. C., Shui C. S., 2016. A multiple type bike repositioning problem. Transport. Res. Part B, 90, 263–278.
Luque B., Lacasa L., Ballesteros F. et al., 2009. Horizontal visibility graphs: exact results for random time series. Phys. Rev. E 80, 593-598.
Lacasa L., Luque B., Ballesteros F., Luque J., Nuo J. C., 2008. From time series to complex networks: the visibility graph. Proc. Nat. Acad. Sci., 105, 4972-4975.
Maria Bordagaray, Luigi dell'Olio, Achille Fonzone, Ángel Ibeas, 2016. Capturing the conditions that introduce systematic variation in bike-sharing travel behavior using data mining techniques. Transport. Res. Part C, 71, 231-248.
Motter A. E., Lai Y. C. Cascade-based attack on complex networks. Physical Review E, 2002, 66(6), 065102.
Newman M E J., 2004. Fast algorithm for detecting community structure in network. Physical Review E, 69, 066133.
Ott E., 2002. Chaos in dynamical systems. Cambridge University Press, Cambridge.
Pal, A., Zhang, Y., 2017. Free-floating bike sharing: Solving real-life large-scale static ebalancing problems. Transport. Res. Part C 80, 92-116.
Pfrommer J., Warrington J., Schildbach G., Morari M., 2014. Dynamic vehicle redistribution and online price incentives in shared mobility systems. IEEE Transactions on Intelligent Transportation Systems, 15, 1567-1578.
Raidl G. R., Hu B., Rainer-Harbach M., Papazek P., 2013. Balancing bike sharing systems: improving a VNS by efficiently determining optimal loading operations. Hybrid Metaheurist. 7919, 130-143.
Rainer H. M., Papazek P., Raidl G. R., Hu B., Kloimüllner C., 2015. PILOT, GRASP, and VNS approaches for the static balancing of bike sharing systems. J. Global Optim., 1-33.
Raviv T., Kolka O., 2013. Optimal inventory management of a bike-sharing station. IIE Trans. 45, 1077-1093.
Raviv T., Tzur M., Forma I. A., 2013. Static repositioning in a bike-sharing system: models and solution approaches. Eur. J. Transp. Logist. 2, 187-229.
Regue R., Recker W., 2014. Proactive vehicle routing with inferred demand to solve the bike sharing rebalancing problem. Transport. Res. Part E, 72, 192-209.
Schuijbroek J., Hampshire R., Hoeve W. J., 2017. Inventory rebalancing and vehicle routing in bike sharing systems. Eur. J. Oper. Res. 257, 992-1004.
Shannon, C. E., 1948. A mathematical theory of communication. Bell System Technical Journal. 27, 379-423.
Shui C. S., Szeto W. Y., 2014. Solving a static repositioning problem in bike-sharing systems using iterated tabu search. Transport. Res. Part E, 69, 180-198.
Shui C. S., Szeto W. Y., 2018. Dynamic green bike repositioning problem-A hybrid rolling horizon artificial bee colony algorithm approach. Trans. Res. D., 60, 119-136.

Szeto W. Y., Liu Y., Ho S. C., 2016. Chemical reaction optimization for solving a static bike repositioning problem. Transport. Res. Part D, 47, 104-135.

Tang Y., Yi N., Mao J. H., 2017. Derived network based on directed limited penetrable visibility graph for time series. J. Syst. Eng., 32, 156-162.

Wang M. G., Chen Y., Tian L. X., Jiang S. M., Tian Z. H., Du R. J., 2016. Fluctuation behavior analysis of international crude oil and gasoline price based on complex network perspective. Appl. Energy 175, 109-127.

Warrington J., Ruchti D., 2019. Two-stage stochastic approximation for dynamic rebalancing of shared mobility systems. Transport. Res. Part C, 104, 110-134.

Wei X. Y., Luo S. D., Nie Y. M., 2019. Diffusion behavior in a docked bike-sharing system. Transport. Res. Part C, 107. 510-524.

Yang Y., Liu Y., Zhou M., Li F., Sun C., 2015. Robustness assessment of urban rail transit based on complex network theory: A case study of the Beijing Subway, Safety Sci. 79, 149-162.

You P. S., 2019. A two-phase heuristic approach to the bike repositioning problem. Applied Mathematical Modelling, 73,651-667.

Zhang D., Yu C. H., Desai J., Lau H. Y. K., 2017. A time-space network flow approach to dynamic repositioning in bike sharing systems. Transport. Res. Part B, 103, 188-207.

Zhang Y., Thomas T., Brussel M. J. G., Maarseveen M. F. A. M., 2016. Expanding Bike-Sharing Systems: Lessons Learnt from an Analysis of Usage. PLOS ONE, 11, 1-25.

Zhang J., Meng M., Wang Z. W. D., 2019. A dynamic pricing scheme with negative prices in dockless bike sharing systems. Transport. Res. Part B, 127, 201-224.

Zhao J. B., Wang J., Deng W., 2015. Exploring bikes haring travel time and trip chain by gender and day of the week. Transport. Res. Part C, 58, 251-264.

Zhou Y., Donner R. V., Marwan N., Donges J. F., Kurths J., 2018. Complex network approaches to nonlinear time series analysis. Physics Reports., 787,1-97.