

Article

Fleet Size and Rebalancing Analysis of Dockless Bike-Sharing Stations Based on Markov Chain

Yong Zhai, Jin Liu *, Juan Du and Hao Wu

National Geomatics Center of China, Beijing 100830, China

* Correspondence: liujin@ngcc.cn

Received: 29 April 2019; Accepted: 25 July 2019; Published: 29 July 2019



Abstract: In order to improve the dynamic optimization of fleet size and standardized management of dockless bike-sharing, this paper focuses on using the Markov stochastic process and linear programming method to solve the problem of bike-sharing fleet size and rebalancing. Based on the analysis of characters of bike-sharing, which are irreducible, aperiodic and positive-recurrence, we prove that the probability limits the state (steady-state) of bike-sharing Markov chain only exists and is independent of the initial probability distribution. Then a new “Markov chain dockless bike-sharing fleet size solution” algorithm is proposed. The process includes three parts. Firstly, the irreducibility of the bike-sharing transition probability matrix is analyzed. Secondly, the rank-one updating method is used to construct the transition probability random prime matrix. Finally, an iterative method for solving the steady-state probability vector is therefore given and the convergence speed of the method is analyzed. Furthermore, we discuss the dynamic solution of the bike-sharing steady-state fleet size according to the time period, so as improving the practicality of the algorithm. To verify the efficiency of this algorithm, we adopt the linear programming method for bicycle rebalancing analysis. Experiment results show that the algorithm could be used to solve the disordered deployment of dockless bike-sharing.

Keywords: Markov chain; steady-state fleet size; rebalancing; transition probability random prime matrix; linear programming

1. Introduction

1.1. Background

The concept of “bike-sharing” [1–4] was proposed the first time in Amsterdam in 1965. Until 2010, Social Bicycles Inc. (SoBi) [5] in the US created a bicycle operation mode by using mobile APP and GPS to realize a quick rent and return of the bicycle, the prototype of the dockless bike-sharing was established [6].

The application of bike-sharing started lately but developed rapidly in China. Taking Hangzhou as an example, a sizeable docked public bicycle system has been built there since 2008. Statistics showed that the scale was increased from 2800 vehicles and 61 stations to 66,000 vehicles and 2411 stations, and the average usage rate of bike-sharing vehicles is about 250,000 times per day. Until 2011, it has become the world’s largest public bicycle system [7–9]. At the same time, urban public bicycles have developed rapidly in various parts of China, but the penetration was not so well. In 2015, OFO Inc. firstly deployed dockless bike-sharing on the university campus, which made the rise of dockless bike-sharing in China. Despite the fact that the development period of the dockless bike-sharing is not so long, but both the development speed and the scale have seen remarkable growth with the encouragement of capital markets and increase of social requirements. By the end of 2017, the number of bike-sharing has exceeded 11 million in China.

The bike-sharing deployed by merchants since 2016 are mainly dockless bicycles [10]. Compared with traditional docked bicycles, it could allow users to pick up and drop off bikes freely instead of at the designated station [11]. The dockless bike-sharing adopts the operation mode of “smart dockless”, “GPS positioning”, “network convenient payment” and “requesting on demand”. It has become a popular way for short-distance green travel in China.

The emergence of dockless bike-sharing is an innovative model of the sharing economy in the Internet era, which facilitates the daily short-distance travel of the public [12]. However, on the other hand, the disordered deployment of dockless bike-sharing has become an increasingly serious problem, because some bike-sharing companies try to keep the competition by blind expansions. The problems of fleet size and rebalancing [13–15] of the dockless bike-sharing stations are mainly illustrated as follows:

- There are too many bicycles in some stations, but the problems of idleness are prominent, which even hinders the passage of pedestrians and cars. In contrast, the number of bicycles in some other stations is not enough to meet the needs of users.
- The response of the bike-sharing replenishment or repositioning transport is not so efficiently, and the rationality of the rebalancing is still a concern.

Therefore, the fleet size and rebalancing of dockless bike-sharing stations are in urgent need of scientific analysis. It is necessary to propose an algorithm to determine the number of shared bicycles rationally and quickly at each station to achieve the goal of standardized management.

1.2. Literature Review

Since 2010, researchers have made tremendous efforts into the research and analysis on sharing policies, operating methods, business models, user behaviors and preferences, station location [16,17], and vehicle rebalancing [18–20] of docked bike-sharing. There exists obvious differences such as the spatial distribution mode [21], the riding route [21], the commute correlation [21], the lease and usage mode, and the charging mode between docked and dockless bike-sharing. More flexibility, convenience and randomness, larger user group, higher frequency of use and wider coverage, determine the different research methodologies on the fleet size and rebalancing of dockless bike-sharing. There are a few researches on dockless bike-sharing as the operation time is relatively short.

In 2017, based on Mobike’s [22] case of the bike-sharing application, Xiong Shenyuhang proposed the deductive method to analyze the impact of China’s Internet-based short-distance travel schemes [23], including the public daily life, government and society. In daily life, dockless bike-sharing has changed the way of life of the public and alleviated the traffic congestion problem. At the government level, dockless bike-sharing is a useful supplement to the government-funded docked bike-sharing, which can be supported by government preferential policies. At the society level, dockless bike-sharing has significantly improved the efficiency of urban travel and reduced urban environmental pollution.

Qi Luo et al. [24] designed the site selection and travel assignment algorithm based on integer programming techniques, composed a multi-mode system integrating the dockless bike-sharing and network carpooling service, and the feasibility of the system is demonstrated. The research was based on the hypothesis that the user has the willingness to ride to the nearest carpool hub to select the carpool service and the required bicycle and carpool vehicles exist. The research used the travel and site data of the taxi and dockless bike-sharing in Manhattan, New York City.

In 2018, based on Mobike’s dockless bike-sharing riding data, Wencong Cui [25] selected Beijing as the research area to study the relationship between the use efficiency of dockless bike-sharing and the urban building density by using a case analysis and big data method. The density of the city is negatively correlated with the bicycle idle time and is positively related to the rebalancing. The authors used the SPSS and ArcGIS software for statistical and spatial analysis.

Ling Pan et al. [26] believe that the rebalancing problem has a significant impact on the service quality and revenue of the dockless bike-sharing company. In order to solve the rebalancing problem,

constructing the model based on the Markov decision process by combining the spatial elements, and the Hierarchical Reinforcement Pricing (HRP) algorithm is developed based on the depth deterministic strategy gradient algorithm to encourage users to rebalance the dockless bike-sharing. Thus, the deep learning framework was constructed. The HRP algorithm used the divide-and-conquer method to find the dependence relationship between time and space elements, which overcame the shortcomings of other algorithms that often ignore spatial information and rely on accurate prediction. The authors believed that the HRP algorithm can improve the quality of the dockless bike-sharing service and optimize bicycle deployment.

In addition, other researchers have published some articles on the transportation value and development path [27], the profit path [28], the rules of behavior [29] and the environment design of crime prevention [30].

1.3. Research Objectives

The fleet size and rebalancing analysis of dockless bike-sharing stations in urban areas is not yet mature. On the one hand, there is a few published articles on this topic, especially the analysis of the theoretical value of the fleet size of bicycle stations is still unrevealed. On the other hand, some articles owe potential theoretical defects, such as the assumption of user incentive. During times of heavy usage, the actual research found that the effect of user incentive mechanism (actually the incentive amount is often small) is not obvious because it is time-consuming. The incentive and punishment mechanism should be used simultaneously as supplementary methods for bike-sharing rebalancing.

The use of dockless bike-sharing is actually a stochastic process. Under the given fleet size of urban dockless bike-sharing, it is necessary to find the principles of the bicycle riding transfer between stations and study the optimal fleet size of bike-sharing stations. Following the lowest rebalancing cost principle, the rebalancing algorithm is studied by using the linear programming method. This is the original motivation of this paper.

In this paper, the public area that deployed dockless bike-sharing is defined as a rental station. In order to facilitate the computing process, a city could be divided into equal squares, which are called grid cells showed in Figure 1 (excluding residential areas, government agencies, parks, water systems, roads, etc.). Each of them can be regarded as a station for dockless bike-sharing pick up and drop off. The fleet size of the deployed dockless bike-sharing stations refers to the number of bicycles in the bike-sharing stations. The total fleet size of bike-sharing refers to the total number of shared bicycles in a city. At the same time, to avoid the interference of major economic activities, bad weather, and users' personal preferences for a different dockless bike-sharing brand, the fleet size and rebalancing of dockless bike-sharing stations from the same company are discussed.



Figure 1. An example region is divided into squares with an equal area.

This paper is an extension of the work originally presented in 2018 ISPRS TC IV Mid-term Symposium “3D Spatial Information Science – The Engine of Change” [31]. Aiming at dynamical optimization and standardized management for the fleet size of dockless bike-sharing, this paper proposes the “Markov chain dockless bike-sharing fleet size solution” algorithm. The structure of this paper could be organized as follows. Firstly, we introduce the background, literature review and research object in Section 1. Secondly, the applicability of Markov chain to dockless bike-sharing is analyzed in Section 2. Thirdly, the methods of solving the fleet size and rebalancing analysis of dockless bike-sharing is discussed in Section 3. Based on the above, the experiment and verification of the proposed algorithm is introduced in Section 4. Finally, the preliminary progress and prospective achievement are discussed in Section 5.

2. Applicability Analysis of Markov chain

2.1. Brief Concept of the Markov Chain

A discrete-state, discrete-time Markov stochastic process is called a Markov chain, named after the Russian mathematician Andrey Markov. Markov chains are widely used in economics, traffic management, game theory, queuing theory, genetics and finance.

A Markov chain is a stochastic model describing a sequence of possible events in which the probability of each event depends only on the state attained in the previous event. We have a set of states, $X = \{X_1, X_2, \dots, X_r\}$. The process starts in one of these states and moves successively from one state to another. Each move is called a step. If the chain is currently in state X_i , then it moves to state X_j at the next step with a probability denoted by P_{ij} , and this probability does not depend upon the states the chain was in before the current state. The probabilities P_{ij} are called transition probabilities. The process can remain in the state it is in, and this occurs with probability P_{ii} . Then,

$$P\{X_{n+1} = j | X_0, \dots, X_n\} = P\{X_{n+1} = j | X_n\} \quad (1)$$

A matrix consisting of one-step transition probabilities of multiple states is called a transition probability matrix, in which the value showed in each of these rows is the probability that the state represented by the diagonal element is transferred to another state, respectively, and the probability that each column is transferred to the state represented by the diagonal element, respectively.

Let P be a matrix of one-step transition probabilities, state space $I = \{1, 2, \dots\}$, then

$$P = \begin{pmatrix} P_{11} & P_{12} & \dots & P_{1n} & \dots \\ P_{21} & P_{22} & \dots & P_{2n} & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} \quad (2)$$

P is called a one-step transition probability matrix of Markov chains, and $P_{ij} \geq 0, i, j \in I, \sum P_{ij} = 1$.

2.2. Markov Chain Characters of Bike-Sharing

Bike-sharing provides people with a convenient way for commuting, and solves the “first/last mile” travel problem (DeMaio, 2009). Based on the bike-sharing usage model, the users mainly rent and return the bicycles between adjacent stations. Of course, the concept of “being adjacent” is relative. There is a situation where users rent and return bicycles between stations that are far away from each other. This section theoretically assumes that there is connectivity or bicycle state transition relationships between any two stations to facilitate analysis.

Assuming that a user rents the shared bicycle from the station j and returns the rented bicycle at the station k within the time period $t_0 - t_1$, regardless of the state of other stations in the same interval, the state of the station k during the transfer is only related to the station j , and independent of the states of other stations before the time t_0 . According to the Markov Chain property, “the past is conditionally independent of the future given the present state of the process” or “given the present

state, the past contains no additional information on the future evolution of the system.” So, the state relationship of shared bicycle stations is in accordance with the basic conditions of the Markov stochastic process application.

Suppose a city has a total of n dockless bike-sharing stations, the states of station j during the time period $t_0 - t_1$ include:

- The user at the station j rents a bicycle ride to any other stations that exists $n - 1$ kind of possible (mutually exclusive), each of which has two states, (e.g., the bicycle at station j and station k). Set the bicycle move from station j to station k , and the two states are denoted as X_j and X_k , respectively. The transition probability of this bicycle process is denoted as $P_{jk}, P_{jk} = P\{X_k|X_j\}$, $k \neq j$.
- The bicycle at station j is not rented or still returned to the station j after being rented, the state is denoted as X_0 , and the transition probability is denoted as P_{jj} .

The bicycle transfer state of station j in the time period $t_0 - t_1$ is showed in Figure 2.

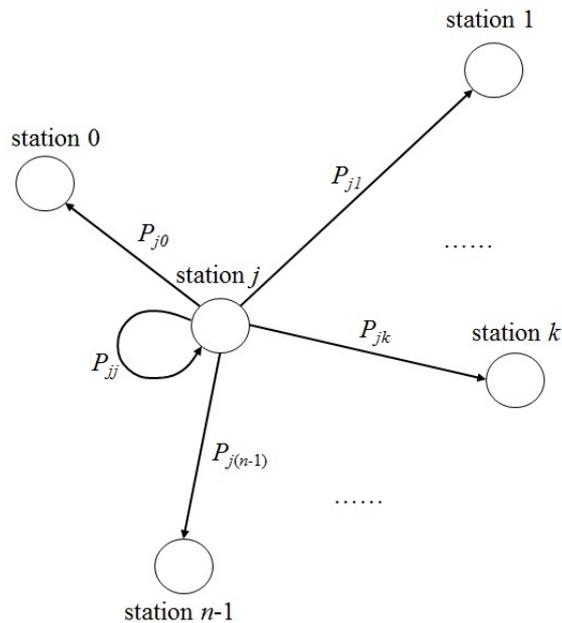


Figure 2. Transition states of station j .

For the probability of dockless bike-sharing discussed in the paper, P_{jk} can be regarded as the ratio of the average number of bicycles that the user rides from station j to station k in the multiple time periods to the average number of bicycles in station j at t_0 . P_{jj} can be regarded as the ratio of the average number of bicycles that are not rented or still returned to station j in the multiple time periods to the average number of bicycles in station j at t_0 .

In summary, the probability P_{jk} and P_{jj} are accumulated to 1.

$$\sum_{k=0}^{n-1} P_{jk} = P_{j0} + P_{j1} + \dots + P_{jj} + \dots + P_{j(n-1)} = 1 \tag{3}$$

$$P_{jk} > 0 \text{ (To ensure connectivity between stations, } P_{jk} \neq 0), k = 0, 1, 2, \dots, n - 1.$$

By analogy, the transition state of all n stations is showed in Figure 3.

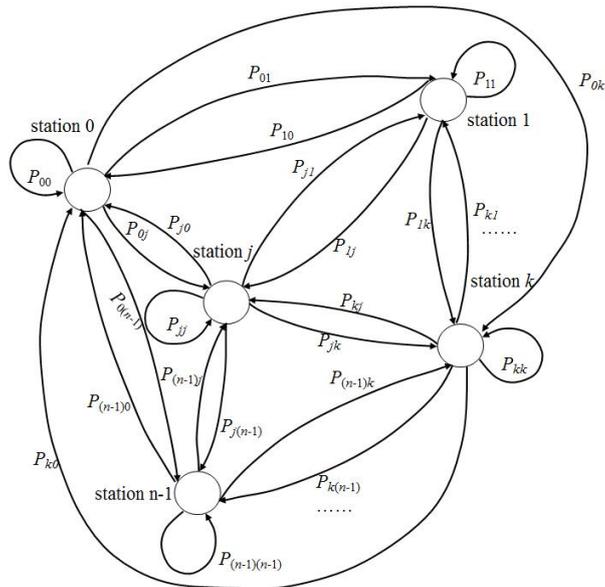


Figure 3. Transition states of all stations.

The transition probability matrix of all n stations is constructed as

$$\mathbf{P} = \begin{pmatrix} P_{00} & P_{01} & \cdots & P_{0j} & \cdots & P_{0(n-1)} \\ P_{10} & P_{11} & \cdots & P_{1j} & \cdots & P_{1(n-1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ P_{j0} & P_{j1} & \cdots & P_{jj} & \cdots & P_{j(n-1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ P_{(n-1)0} & P_{(n-1)1} & \cdots & P_{(n-1)j} & \cdots & P_{(n-1)(n-1)} \end{pmatrix} \tag{4}$$

$$P_{jj} = 1 - \sum_{k=0; k \neq j}^{n-1} P_{jk}, j \in [0, n-1]$$

Since all the elements in the transition probability matrix \mathbf{P} are nonnegative and the sums of elements in any row are all 1, \mathbf{P} is a random matrix.

At a certain time, the ratio of the number of bicycles at each station to the total number of bicycles deployed in the city is called the probability of the station, and the probability values of all stations constitute the probability vector at this time. It is assumed that the initial probability vector of the dockless bike-sharing stations at time t_0 is $\pi^{(0)}$, and the probability vector at time t_1 is $\pi^{(1)}$. The bike-sharing transition probability matrix in the time period $t_0 - t_1$ is \mathbf{P} , by the Markov chain property, it can be obtained:

$$\pi^{(1)} = \pi^{(0)} \mathbf{P} \tag{5}$$

The t_0 moment can be regarded as the 0th transfer of the bike-sharing Markov chain, and the t_1 moment can be regarded as the 1th transfer. With the passage of time, the probability vectors after the $k - 1$ steps and k steps transfer are $\pi^{(k-1)}$ and $\pi^{(k)}$, respectively. The relationship between the two is

$$\pi^{(k)} = \pi^{(k-1)} \mathbf{P} \ (k \geq 1) \tag{6}$$

Iterate over Equation (1) and one can get

$$\pi^{(k)} = \pi^{(k-1)} \mathbf{P} = \pi^{(k-2)} \mathbf{P}^2 = \cdots = \pi^{(0)} \mathbf{P}^k \ (k \geq 1) \tag{7}$$

2.3. Proof of Steady-State Existing in Bike-Sharing Markov Chain

The state of the dockless bike-sharing Markov chain based on a city is limited and has the following properties:

- Irreducible. Let C be a non-empty subset of the state space. If $i \in C$ and $k \notin C$, the connectivity probability $P_{ik} = 0$, C is called a closed set. If all the states in C are interconnected, C is called an irreducible closed set [32]. If Markov chain's state space is an irreducible closed set, it is called irreducible. The total number of bike-sharing stations in a city is limited, all belong to the same category and theoretically there is a bidirectional connectivity relationship between any two stations (except for prohibited areas such as residential areas, government agencies, parks, water systems, and roads, etc.). The bike-sharing Markov chain theoretically has irreducible properties [33].
- Aperiodic. Each state in the Markov chain can only be accessed at periodic intervals, then it is periodic, otherwise it is aperiodic. Theoretically, if the irreducible transition probability matrix is a prime matrix, the Markov chain is aperiodic. Theoretically, if the irreducible transition probability matrix P is a prime matrix (there is only one non-negative irreducible matrix with only one eigenvalue falling on its spectral circle) [34] the Markov chain is aperiodic. According to Carl D. Meyer's research, if a random matrix is irreducible and has at least one positive diagonal element, it is a prime matrix (sufficient condition) [35]. Investigate the transition probability matrix P , which is an irreducible nonnegative square matrix with n diagonal element P_{jj} . In the actual rental of dockless bike-sharing, if all the diagonal elements of P are 0, which are the situations where all bicycles at all stations in the time period $t_0 - t_1$ are rented by the users or there is no return of bicycles to the original stations, this is inconsistent with a real-life situation. Observing the use of bicycles, it can be concluded that the diagonal elements with no internal zero are more prevalent, that is the trace of the matrix P is not 0 ($tr(P) \neq 0$). Therefore, the dockless bike-sharing Markov chain is aperiodic (non-periodic).
- Positive-recurrence. If and only if starting from state j , the random process can eventually return to state j with probability 1, then state j is considered as a recurrence [32]. When the average return time of state j is finite, it is called positive-recurrence. Liu Cihua's argument in the "Random Process" (Fourth Edition) has shown that irreducible finite Markov chains must be recurrence [32]. It has been discussed above that the bike-sharing Markov chain has the properties of being finite and irreducible, so it is a positive-recurrence.

It can be seen from the above that the dockless bike-sharing Markov chain has the above three properties, and the following theorem is established:

Theorem 1. For any irreducible, aperiodic and positive recurrence Markov chain, the limit state probability exists only and independent of the initial probability distribution. This limit state is called steady-state and the limit state probability is called steady-state probability.

When k steps of the bike-sharing Markov chain transition are big enough, one has the following,

$$\lim_{k \rightarrow \infty} \pi^{(k)} = \pi \quad (8)$$

where

$\pi^{(k)}$ = the probability vector after k steps transfer

π = a steady-state probability vector

From formula (2), one has the following,

$$\lim_{k \rightarrow \infty} \pi^{(k)} = \lim_{k \rightarrow \infty} (\pi^{(0)} P^k) = \pi \quad (9)$$

Therefore, the main idea of the algorithm in the paper is based on multiplying the initial probability vector of the dockless bike-sharing stations by the higher power of the transition probability matrix, that is $\pi^{(0)}P^k$. The computed vector will approximate the steady-state probability vector with the increase of power. With the increase of iteration, the probability vector of the dockless bike-sharing stations will gradually stabilize. Each element of the steady-state probability vector is the steady-state probability of bicycles at each bike-sharing station, and multiplying them by the total number of shared bicycles in the city respectively, one obtains the steady-state fleet size of the bike-sharing at each station. Solving the steady-state probability vector of the bike-sharing Markov chain can optimize the configuration of the fleet size for each dockless bike-sharing station. The above iterative method is called the power method.

3. Methodology for Fleet Size and Rebalancing of Bike-Sharing

To solve the steady-state probability vector of dockless bike-sharing stations, we must first determine all the elements of the transition probability matrix P . We need to observe and count the state transitions of n stations in the city during the time period $t_0 - t_1$. For every station, it is necessary to count the specific destination station for each user riding out of the bicycle, or for every station, it is necessary to count the specific source station that each user rides back of the bicycle.

The relationship between n stations is n^2 . In addition, it is also necessary to compute the initial value of the fleet size of each station in the same segment, so n stations have n numbers in total. And one should consider the total fleet size of bicycles deployed by the bike-sharing company in the city. The total number of statistics to be counted is not less than $n^2 + n + 1$.

On the basis of the usage statistics of the bike-sharing and the computation of the transition probability, the transition probability matrix of the bike-sharing (sparse matrix) is constructed first. Second, the reducibility of the transition probability matrix is analyzed. If it is irreducible, the steady-state probability is directly obtained by the power method; if it is reducible, the rank-one updating method is used to compute the steady-state probability. Combined with the practical application of bike-sharing, the algorithm is improved to solve the steady-state fleet size of bicycles at different time periods in one day. On this basis, the linear programming method is used for the bicycle rebalancing analysis. The Fleet size and rebalancing analysis of dockless bike-sharing stations framework is showed in Figure 4.

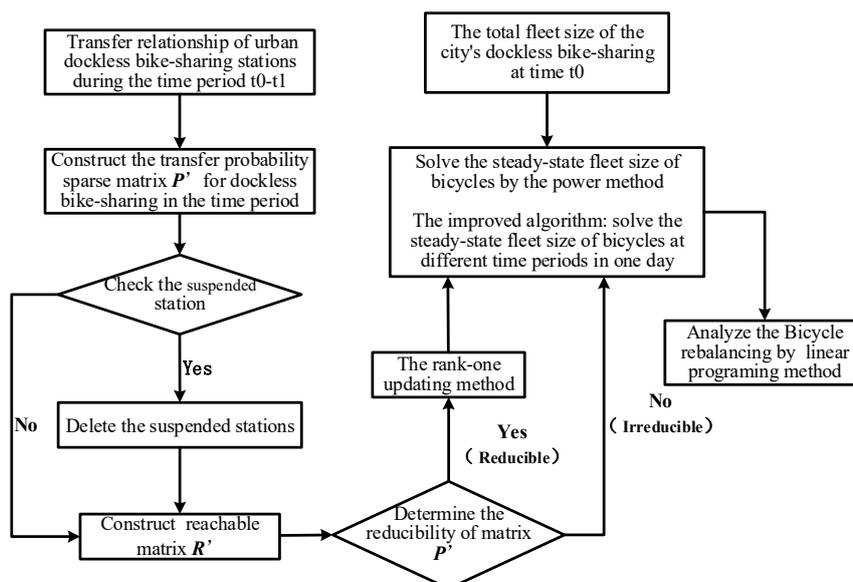


Figure 4. Fleet size and rebalancing analysis of dockless bike-sharing stations framework.

3.1. Analysis of Fleet Size

3.1.1. Sparse Matrix Construction and Reducibility Determination

Based on the observation of the usage of dockless bike-sharing, it is found that for any station j , the riding of the shared bicycle is mainly performed at its adjacent stations. The number of ride-in and ride-out bicycles between the stations that are far away is very small. The user's single riding distance statistics are showed in Table 1 [36].

Table 1. Bike-sharing user single ride distance statistics.

Serial Number	Single Ride Distance (d)	The Proportion of Users
1	$d \leq 3$ km	41.60%
2	$3 < d \leq 5$ km	39.40%
3	$5 < d \leq 10$ km	15.20%
4	$d > 10$ km	3.80%
	Sum	100.00%

According to Table 1, the bike-sharing stations can be divided into the following three categories.

- Close stations. For any station j , another station within 10 km away has a more frequent bike-sharing relationship with it. This paper refers to this kind of station as the close station of station j .
- Peripheral stations. For any station j , the station that is more than 10 km away is a peripheral station. According to Table 1, the average value of the peripheral stations to the total number of stations could be set to 4%.
- Unrelated stations. This paper refers to a station with no bicycle transfer relationship between station j as an unrelated station. The transition probability between the unrelated station and the station j is set to zero. An unrelated station with a bicycle initial value of zero is called a suspended station.

Based on the above analysis, a new dockless bike-sharing transition probability matrix P' is constructed, which is a sparse matrix obviously. Set each station have an average number of close and peripheral stations are \hat{m} , then there are about $\hat{m} \times n$ non-zero elements in P' . For medium and large cities, $\hat{m} \ll n$, $\hat{m} \times n \ll n^2$.

This method can be referred to as the "Markov chain dockless bike-sharing fleet size solution". Using this method, one can deduce that the mutual riding relationship between n stations is reduced to $\hat{m} \times n$. In addition, it is also necessary to count the initial fleet size of each station, as n stations have a total of n values. If the total fleet size of bicycles is deployed by the bike-sharing company in the city, the total number of statistics to be counted is not less than $\hat{m} \times n + n + 1$, which is significantly smaller than the traditional total number of statistics $n^2 + n + 1$.

After the transition probability matrix is sparse, the station interoperability is affected due to the unreachability of some stations. The irreducible nature of the transition probability matrix may be changed to be reducible. The steady-state condition of the matrix is destroyed and the steady-state fleet size cannot be solved. Therefore, the suspended station check is firstly performed on the sparse matrix P' . If the elements in the row and column of the P' diagonal element a_{jj} are both zero (including $a_{jj} = 0$), the rows and columns are deleted from P' .

On this basis, the matrix P' is checked for reducibility. If P' is irreducible, the power method is directly performed (see 3.1.3); otherwise, the rank-one updating method of the sparse matrix needs to be performed first to satisfy the irreducible condition, and then the power method is carried out. The following method is used to determine the reducibility of the matrix P' :

Construct the reachable matrix R' of the matrix P' by replacing all non-zero elements in P' with one and unchanging zero elements. The determination of matrix reducibility is equivalent to the determination of the connectivity of the reachable matrix.

Determining the connectivity between any two stations j and k cannot be argued only based on whether the element a_{jk} in R' is equal to zero. This only indicates that the direct connection between the two stations does not exist, and it is necessary to verify the possibility of connectivity through other stations.

Arguing station j and station k are connected by w ($1 \leq w \leq n-2$, n is the total number of stations) stations, one needs to consider the arrangement of w stations. That is, one needs to compute the possible arrangement of A_{n-2}^w . Analyzing whether the multiplications of the corresponding elements in each case of the arrangement is equal to zero, if there is a case that is not equal to zero, stations j and k are connected. If there exists a case where it is equal to zero, stations j and k are not connected. If any two stations in the matrix R' have a connectivity relationship, the sparse matrix P' is irreducible. If there is a non-connective relationship between the two stations, P' is reducible.

One can argue that station j and station k are connected by w stations, the maximum number of connection paths is $A_{n-2}^1 + A_{n-2}^2 + \dots + A_{n-2}^{n-2}$, $n \geq 3$.

The computation complexity of the connectivity argument is computed as follows (See proof of the below equation in Appendix A):

$$\begin{aligned} O(n) + A_{n-2}^0 &= A_{n-2}^0 + 2A_{n-2}^1 + 3A_{n-2}^2 + \dots + (n-1)A_{n-2}^{n-2} \\ &= \binom{n-2}{0} + 2! \binom{n-2}{1} + 3! \binom{n-2}{2} + \dots + (n-1)! \binom{n-2}{n-2} \\ &= \sum_{k=0}^{n-2} (k+1)! \binom{n-2}{k} \\ &= (n-2)! \sum_{k=0}^{n-2} \frac{k+1}{(n-2-k)!} \\ &= (n-2)! \sum_{k=0}^{n-2} \frac{n-k-1}{k!} \\ &\geq (n-2)! \sum_{k=0}^{n-2} \frac{n}{k!} = \frac{n}{2} (n-2)! \sum_{k=0}^{n-2} \frac{1}{k!} \end{aligned}$$

Through Taylor's formula, one obtains,

$$\sum_{k=0}^{n-2} \frac{1}{k!} \approx e$$

Then

$$O(n) \geq \frac{n}{2} (n-2)! e - 1$$

The above is only the complexity of the connectivity computation between one station pairs. Considering that the maximum number of different station pairs in the matrix P is the $n^2 - n$, assuming that m ($1 \leq m \leq n^2 - n$) station pairs are directly connected, the total connectivity computation complexity is not less than $(n^2 - n - m) \left(\frac{n}{2} (n-2)! e - 1 \right)$. R. Tarjan's strongly connected components algorithm [37] significantly reduces the computational complexity of the matrix reducibility argument to $O(n)$. In practice, we can use software tools such as MATLAB to argue the reducibility of the matrix.

3.1.2. The Rank-One Updating of the Reducible Matrix

The rank-one updating method first appropriately reduces the probability of non-zero elements in the transition probability sparse matrix P . Then, the sum of the transitional probabilities of peripheral stations (not exceeding 4%) is evenly distributed to all stations. Under the premise that the sum of the elements of each row of the matrix is one, a new random matrix that satisfies the solving of the steady-state probability of dockless bike-sharing is constructed.

In the specific rank-one updating operation, firstly, P' can be multiplied by one adjustment parameter θ (that is $\theta P'$, $0 < \theta < 1$, and θ is close to one) to decrease the non-zero elements appropriately. Then the sum of the elements of each row in the matrix P is θ . Next, each element in P adds $\frac{1-\theta}{n}$ separately, and the sum of the increments in each row is $(\frac{1-\theta}{n}) \times n = 1 - \theta$, so the sum of the elements of each row is $\theta + (1 - \theta) = 1$. Therefore, the sum of the elements of each row of the matrix remains unchanged. From the above, a new transition probability matrix C is constructed. The above operation can be expressed as

$$C = \theta P' + \frac{1-\theta}{n} ee^T \tag{10}$$

e^T is the transposition for the unit column vector e . The rank of the matrix formed by $(\frac{1-\theta}{n})ee^T$ is one, so the above operation is called the rank-one updating for P . According to Table 1, about 96% of the users have a single riding distance within 10 km; thus the θ in the C can be selected as 0.96, and the computation result of the rank-one updating part is $n \times n$ square matrix. One has the following,

$$\frac{1-\theta}{n} ee^T = \frac{0.04}{n} ee^T = \begin{pmatrix} \frac{0.04}{n} & \dots & \frac{0.04}{n} \\ \vdots & \ddots & \vdots \\ \frac{0.04}{n} & \dots & \frac{0.04}{n} \end{pmatrix} \tag{11}$$

In summary, we know that C is a random prime matrix, which satisfies the irreducible, aperiodic and positive-recurrence property of the dockless bike-sharing Markov chain. The steady-state condition of the transition probability matrix is established.

3.1.3. Steady-State Fleet Size Solution

When computing the steady-state probability vector of C by the power method, one has

$$\begin{aligned} \pi^{(k)} &= \pi^{(k-1)} C = \pi^{(k-1)} \left(\theta P' + \frac{1-\theta}{n} ee^T \right) \\ &\because \pi^{(k-1)} e = 1 \\ \therefore \pi^{(k)} &= \theta \pi^{(k-1)} P' + (1-\theta) \frac{1}{n} e^T \end{aligned} \tag{12}$$

Although the matrix C is a dense matrix, it can be known from Equation (6) that we actually perform vector-matrix operations on the original sparse matrix P' when performing the power method, and the correlation operation of the dense matrix is not performed. The change in the element after the rank-one updating was not directly involved in the computation, therefore the time complexity of this algorithm is still $O(n)$.

Continue iterating on the $\pi^{(k)}$, one can obtain,

$$\begin{aligned} \pi^{(k)} &= \theta \pi^{(k-1)} P' + (1-\theta) \frac{1}{n} e^T \\ &= \theta \left(\theta \pi^{(k-2)} P' + (1-\theta) \frac{1}{n} e^T \right) P' + (1-\theta) \frac{1}{n} e^T \\ &= \theta^2 \pi^{(k-2)} P'^{(2)} + \theta(1-\theta) \frac{1}{n} e^T P' + (1-\theta) \frac{1}{n} e^T \\ &\vdots \\ &= \theta^k \pi^{(0)} P'^{(k)} + (1-\theta) \frac{1}{n} e^T \left(\theta^{k-1} P'^{(k-1)} + \theta^{k-2} P'^{(k-2)} + \dots + \theta P' + 1 \right) \end{aligned} \tag{13}$$

Substituting $\theta = 0.96$ into the above equation yields,

$$\pi^{(k)} = 0.96^k \pi^{(0)} P'^{(k)} + 0.04 \frac{1}{n} e^T \left(0.96^{k-1} P'^{(k-1)} + 0.96^{k-2} P'^{(k-2)} + \dots + 0.96 P' + 1 \right) \tag{14}$$

When n is large enough, one can obtain,

$$\pi^{(k)} \approx 0.96^k \pi^{(0)} \mathbf{P}'^{(k)} \quad (15)$$

After the rank-one updating, the probability value of the non-zero element P_{jk} in the sparse matrix \mathbf{P}' is changed as follows:

$$\Delta P = \left(P_{jk} \theta + (1 - \theta) \frac{1}{n} \right) - P_{jk} = \left(0.96 P_{jk} + 0.04 \frac{1}{n} \right) - P_{jk} = 0.04 \left(\frac{1}{n} - P_{jk} \right) \quad (16)$$

If the number of bike-sharing stations in the city is large (such as $n \geq 100$), $|\Delta P| \approx 0.04 P_{jk}$, the probability value of non-zero elements changes very little. In addition, after the rank-one updating, the probability value of each zero element is changed to: $\frac{(1 - \theta)}{n} = \frac{0.04}{n}$. If $n \geq 100$ in the city, the corrected probability value of each zero element is no more than 0.0004.

The larger the city size, the higher the number of the bike-sharing station is, and the smaller the impact of the rank-one updating on the \mathbf{P}' element. Therefore, after the rank-one updating of the sparse matrix \mathbf{P}' , the constructed new transition probability matrix \mathbf{C} does not change the main state transition relationship between the bike-sharing stations. For the bike-sharing operation with low accuracy requirements and high response speed, the rationality of the solution results can be guaranteed.

3.1.4. Steady-State Fleet Size Convergence Determination Method

The dockless bike-sharing transition probability matrix \mathbf{P} is a random matrix, the max row sum matrix norm $\|\mathbf{P}\|_{\infty} = \max_{0 \leq i \leq n-1} \sum_{j=0}^{n-1} |P_{ij}| = 1$. According to the relationship between the spectral radius and matrix norm, the spectral radius of the matrix in any complex field is not greater than the max row sum matrix norms, that is $\rho(\mathbf{P}) \leq \|\mathbf{P}\|_{\infty} = 1$. It is also defined that the spectral radius $\rho(\mathbf{P})$ is the largest value of the eigenvalues of the matrix \mathbf{P} (the absolute value of the dominant eigenvalue), so it is known that the absolute value of the dominant eigenvalue of \mathbf{P} is $|\lambda_1| \leq 1$.

At the same time, since \mathbf{P} is a random matrix, $\mathbf{P}e = e = 1 \times e$, that is, 1 is an eigenvalue of \mathbf{P} .

From the above, the dominant eigenvalue of the bike-sharing transition probability matrix \mathbf{P} is $\lambda_1 = 1$. For a random prime matrix, the dominant eigenvalues are unique [38], therefore, without loss of generality, we can make the following $n - 1$ eigenvalues have the following relationship: $|\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|$. Then $1 > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|$.

Let eigenvalues of the random prime matrix \mathbf{C} be $\{\mu_1, \mu_2, \dots, \mu_n\}$, where $\mu_1 = 1$, without loss of generality, one can make the following $n - 1$ eigenvalues have the following relationship: $1 > |\mu_2| \geq |\mu_3| \geq \dots \geq |\mu_n|$. According to the proof of Carl D. Meyer, $\mu_i = \theta \lambda_i$, $i = 2, 3, \dots, n$, λ_i is the eigenvalues of the dockless bike-sharing transition probability sparse prime matrix \mathbf{P} . Then $\mu_2 = \theta \lambda_2$. Since $|\lambda_2| \leq 1$, $|\mu_2| \leq \theta$.

Based on the random prime matrix \mathbf{C} , we compute the steady-state probability vector of the bike-sharing fleet size by the power method, and its convergence rate is $|\mu_2|^k \rightarrow 0$ [34,35]. Since $|\mu_2| \leq \theta < 1$, the convergence rate does not exceed $|\theta|^k \rightarrow 0$. If the precision of the computation is one digit after the decimal point, set $|\theta|^k = 0.1$, since $\theta > 0$, $\theta^k = 0.1$, we can take the base 10 logarithm on both sides of the equation to get $k = (-1) / \log_{10} \theta$. In this paper, we set $\theta = 0.96$, the number of iterations does not exceed $(-1) / \log_{10} 0.96 \approx 56$ times.

If the minimum time period of the bike-sharing state transition is in unit of day, the convergence value of the power method iteration can be regarded as the time when the overall stations of a city reach the steady-state.

Combining the iterative process of the power method, when computing the steady-state fleet size probability vector, the steady-state can be argued by computing the absolute value of the difference between the corresponding elements of the adjacent probability vectors. Set

$$\pi^{(k)} = (x_1, x_2, \dots, x_n) \quad (17)$$

$$\pi^{(k+1)} = (y_1, y_2, \dots, y_n) \quad (18)$$

When every absolute value of the difference is less than a sufficiently small constant α , one obtains,

$$\left(|y_1 - x_1| \leq \alpha\right) \cup \left(|y_2 - x_2| \leq \alpha\right) \cup \dots \cup \left(|y_n - x_n| \leq \alpha\right), \alpha = \frac{1}{S} \quad (19)$$

where S = the total fleet size of dockless shared bicycles in a city.

At this point, the bike-sharing transfer can be considered to have reached the steady-state. For the convenience of computation, set $\delta = \|\pi^{(k+1)} - \pi^{(k)}\|_1 = \sum_{i=1}^n |y_i - x_i|$. When δ is small enough (may also take $1/S$), it can be considered that two adjacent vectors are equal and the dockless bike-sharing Markov chain reaches the steady-state.

3.1.5. Improvement of Bike-Sharing Fleet Size Algorithm

Through the practical application of dockless bike-sharing, it is not difficult to find the following characteristics: On the workdays, the use of dockless bike-sharing is tightly coupled with different time periods every day, that is, it is used frequently during peak hours and afternoons, moderately used at noon, and less used at night. Therefore, the working day time is divided as shown in Table 2.

Table 2. Bike-sharing user single ride distance statistics.

State Cycle	Time Period
T_1	[7:00–9:00]
T_2	[9:00–11:00]
T_3	[11:00–13:00]
T_4	[13:00–17:00]
T_5	[17:00–19:00]
T_6	[19:00–22:00]
T_7	[22:00 to the next morning 7:00]

The 24 h in Table 2 is a change cycle of the bike-sharing station status. For bike-sharing companies, the change of the fleet size of bike-sharing stations in different time periods can provide an important basis for the formulation of the rebalancing strategy. Therefore, it is necessary to further improve the algorithm.

According to the analysis provided in chapter 2, we can see that the changes of the bicycle state in the seven time periods all meet the basic requirements of discrete-time Markov chain applications. Therefore, the algorithm can be further refined to dynamically compute the transition probability matrix of each major time period of the day, and then compute the steady-state fleet size of each bike-sharing station in each time period.

The initial fleet size of bike-sharing of station j at different time periods in a day is S_{T_i} , $i \in (1, 7)$, respectively. Set the initial fleet size of the bike-sharing of station j in each time period of the $k+1$ th day be $S_{T_i}^{(k+1)}$, $i \in (1, 7)$. In the i -th time period, the number of rides out of station j is $G_{T_i}^{(k+1)}$ and the number of rides is $I_{T_i}^{(k+1)}$. Then

$$S_{T_i}^{(k+1)} = S_{T_{(i-1)}}^{(k+1)} + I_{T_i}^{(k+1)} - G_{T_i}^{(k+1)}, i \neq 1 \quad (20)$$

$$S_{T_i}^{(k+1)} = S_{T_6}^{(k)} + I_{T_7}^{(k)} - G_{T_7}^{(k)}, i = 1$$

Knowing the initial value S_{T_i} of bike-sharing for a certain time period in station j , and the number of shared bicycles G_{T_i} from station j to station k , then the bike-sharing transition probability of station j to station k in this time period is $P_{T_i-jk} = G_{T_i}/S_{T_i}$, $P_{T_i-jj} = 1 - \sum_{k=0}^{n-1} P_{T_i-jk}$, $k \neq j$. Set the transition probability matrix at different time period in a day as $P'_i, i \in (1, 7)$, substituting P_{T_i-jk} and P_{T_i-jj} into the P'_i to construct the transition probability matrix.

Since $P'_i (i = 1, 2, \dots, 7)$ is a random matrix, the $\prod_{i=1}^7 P'_i$ is also a random matrix. Set $Q = \prod_{i=1}^7 P'_i$ (Q can be referred to as bike-sharing total transition probability matrix). If Q is irreducible, the power method is directly performed; if Q is reducible, the rank-one updating method is needed and then the power method is performed. The steady-state probability vector π is obtained by power method.

After reaching the steady-state, the fleet size of each bike-sharing station in each time period of the day can be computed separately. The steady-state fleet size vector of bike-sharing stations at the morning peak in a day is equal to the total size of the bicycles multiplied by the steady-state probability vector.

$$S_{T_1} = S \times \pi \quad (21)$$

According to formula (9), the early morning peak fleet size of bike-sharing (the bicycle station fleet size vector at 7:00 in the morning) is obtained. The vector is multiplied by the state transition probability matrix of other consecutive time periods, thus the fleet size vector of other time periods of the day can be obtained.

The fleet size vector of n bike-sharing stations in other time periods $T_i, i \in (2, 7)$ is solved as follows:

$$S_{T_i} = S_{T_{i-1}} \times P'_{i-1} = S_{T_{i-2}} \times P'_{i-2} \times P'_{i-1} = \dots = S_{T_1} \left(\prod_{j=1}^{i-1} P'_j \right) = S \times \pi \left(\prod_{j=1}^{i-1} P'_j \right) \quad (22)$$

3.2. Analysis of Rebalancing

In practical applications, the fleet size decision of each bike-sharing station can be made by referring to the computation results under steady-state conditions, and optimized by considering factors such as the land area of stations. If the steady-state fleet size of the bike-sharing station is greater than the maximum number of bicycles allowed in the actual area, set station j for this type of stations, if P_{jj} is larger (such as $P_{jj} \geq 5\%$), it means that such stations have more idle bicycles, and the bicycles removal adjustment is required. Otherwise, if P_{jj} is within the normal range of values, indicating that the steady-state bicycle fleet size of such stations exceeds the load capacity of the station land area, the dynamic rebalancing of the number of bicycles at the station is required. The above two situations are collectively referred to as the fleet size rebalancing of dockless bike-sharing. The following is an analysis of the bike-sharing fleet size rebalancing optimization based on the linear programming method.

First, computing the actual number of bicycles that can be accommodated based on the available area of each station, and constructing the actual fleet size vector γ of the bicycle station, let the vector $\eta = S \times \pi - \gamma$. Considering the elements in η , set l_1 elements equal to 0, l_2 elements equal to negative numbers, and l_3 elements equal to positive numbers, then $l_1 + l_2 + l_3 = n$.

For the positive elements in η , we construct the bicycle moving-out fleet size vector $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_{l_3})$, $\alpha_j = S \times \pi_j - \gamma_j$, $1 \leq j \leq l_3$; for the negative elements in η , we construct the bicycle moving-in fleet size vector $\beta = (\beta_1, \beta_2, \dots, \beta_{l_2})$, $\beta_k = S \times \pi_k - \gamma_k$, $1 \leq k \leq l_2$. Constructing moving-out stations sequence vector $\omega = (1, 2, \dots, l_3)$, and moving-in stations sequence vector $v = (1, 2, \dots, l_2)$.

Let d_{jk} be the distance between station j and k . For the convenience of computation, let d_{jk} be the straight line distance between the diagonal intersection points of the grids where station j and k are located. The rectangular coordinate system is established with the lower left corner of the city bike-sharing station grid as the origin, and the coordinates of station j and k are (x_j, y_j) , (x_k, y_k) , $d_{jk} = \sqrt{(x_k - x_j)^2 + (y_k - y_j)^2}$.

Let $\xi_{jk}(j \in \omega, k \in \nu)$ denote the number of bicycles that are transferred to the moving-in station k through transport trucks from the moving-out station j . In order to optimize bicycle rebalancing, bicycle service providers use transport trucks to load and unload bicycles between moving-out stations and moving-in stations. The lowest rebalancing cost is the goal pursued by the businesses. For the rebalancing stations, during a rebalancing period (such as noon time), the fewer bicycles are dispatched by the transport trucks, and the shorter the distance traveled, the lower the rebalancing cost. This paper does not directly compute the rebalancing cost, but uses the multiplication of the bicycle number moved out of the stations and the distance traveled by the transport truck to dispatch [39] these bicycles as a rebalancing optimization. Then the solution is as follows:

$$\begin{aligned}
 & \min \sum_{j=1}^{l_3} \sum_{k=1}^{l_2} d_{jk} \xi_{jk} & (23) \\
 & \text{s.t.} \sum_{k=1}^{l_2} \xi_{jk} = \alpha_j \\
 & \sum_{j=1}^{l_3} \xi_{jk} \leq -\beta_k \\
 & \sum_{j=1}^{l_3} \alpha_j \leq \sum_{k=1}^{l_2} |\beta_k| \\
 & d_{jk} > 0, \xi_{jk} \in \mathbb{Z}^+ \\
 & \forall \alpha_j \in \alpha, \forall \beta_k \in \beta \\
 & \alpha_j = S \times \pi_j - \gamma_j, \beta_k = S \times \pi_k - \gamma_k \\
 & \forall j \in \omega, \forall k \in \nu \\
 & \omega = (1, 2, \dots, l_3), \nu = (1, 2, \dots, l_2)
 \end{aligned}$$

Equation (11) is a special type of transportation problem [40], which can be solved by the Lingo software to obtain the optimal solution (lowest cost). Equation (11) is based on the premise that the total number of moving-out bicycles is not greater than the total number of moving-in bicycles, that is $\sum_{j=1}^{l_3} \alpha_j \leq \sum_{k=1}^{l_2} |\beta_k|$. When the total number of moving-out bicycles is greater than the total number of moving-in bicycles, that is $\sum_{j=1}^{l_3} \alpha_j > \sum_{k=1}^{l_2} |\beta_k|$, which means that the number of shared bicycles exceed the capacity of the urban space area, and the number of bicycles needs to be reduced. When $\sum_{j=1}^{l_3} \alpha_j = \sum_{k=1}^{l_2} |\beta_k|$, it can be considered that the ideal state is reached, and at this time Equation (11) has a feasible solution.

The computation results show that the rebalancing optimization is mainly carried out between the moving-out stations and the nearby moving-in stations. In real-life situation, the moving-out stations and the nearby moving-in stations can be unified for rebalancing management. One can use incentives to encourage the users to pick up the bicycles at the moving-out stations, and use the penalty strategy to reduce the number of users who drop off the bicycles at the moving-out stations, as an auxiliary means to further reduce the cost pressure of rebalancing optimization.

3.3. Algorithmization

Algorithm 1: Markov chain dockless bike-sharing fleet size solution

Input: The initial fleet size of the bike-sharing $\pi^{(0)}$ at 7 : 00 am, the transition probability matrix of seven time periods $P'_1, P'_2, P'_3, P'_4, P'_5, P'_6, P'_7$ in one day, the adjustment parameter $\theta(0.96)$, the total fleet size of the urban bicycles S , the actual fleet size vector of bike-sharing stations γ .

Output: The steady-state fleet size of bike-sharing stations during the seven time periods in a day.

```

1:   Build the transition probability matrix  $Q = \prod_{i=1}^7 P'_i$  of dockless bike-sharing stations;
2:   Argue the reducibility of  $Q$ . If it is irreducible, let  $C = Q$ , turn to step 5;
3:   If  $Q$  is reducible, perform rank-one updating on  $Q$  and construct  $C = \theta Q + [(1 - \theta)ee^T]/n$ ;
4:    $i = 1$ ;
5:   FOR{
6:      $\pi^{(i)} = \pi^{(i-1)}C$ ;
7:      $\delta = \|\pi^{(i)} - \pi^{(i-1)}\|_1 = \sum_{i=1}^n |y_i - x_i|$ ;
8:     IF  $\delta \leq \frac{1}{s}$ 
9:        $\pi_{T_1} = \pi^{(i+1)}$ ;
10:    EXIT
11:    ELSE
12:       $i++$ ;
13:    ENDIF
14:    /*Compute the steady-state fleet size probability vector of dockless bike-sharing stations.*/
15:     $j = 2; m = 1$ ;
16:    FOR{
17:       $S_{T_j} = S\pi_{T_1} \prod_{m=1}^{j-1} P'_m$ ;
18:       $j++; m = 1$ ;
19:      IF  $j > 7$  THEN EXIT;
20:    } /*Compute the fleet size of the dockless bike-sharing stations at different time periods in one day
    respectively.*/
21:    Solve the linear programming equation  $\min \sum_{j=1}^{I_3} \sum_{k=1}^{I_2} d_{jk} \xi_{jk}$ , and obtain the optimal moving-in stations
    collection for each moving-out stations in different time periods.

```

4. Experiment and Verification

This section uses commercial data and simulation data to verify the feasibility of Algorithm 1. The commercial data is provided by Mobike Inc. [22], which is more than 1.14 million lines. After the commercial data verification, the technical experts of Mobike Inc. believe that the steady-state fleet size of the bike-sharing solved by Algorithm 1 can be used as the baseline reference data in practical operations. However, due to the commercial confidentiality agreement signed with Mobike Inc., this paper mainly shows the specific process and results of the simulation experiment and verification.

The simulation experiment proves that Algorithm 1 is effective and can provide decision support for bike-sharing deployment operations. The simulation data is generated randomly by R language, and the matrix computation is performed according to Algorithm 1 to illustrate the feasibility of obtaining the steady-state fleet size of the station for decision support. On this basis, by comparing the time required to reach the steady-state under the different total fleet size of bicycles and different number of stations, the rule of the steady-state realization and the effectiveness of Algorithm 1 is further verified. For ease of explanation, the simulation experiment takes 20 stations as an example, and sets the total fleet size of bike-sharing as $S = 500$. The experiment and verification process of the steady-state probability is as follows,

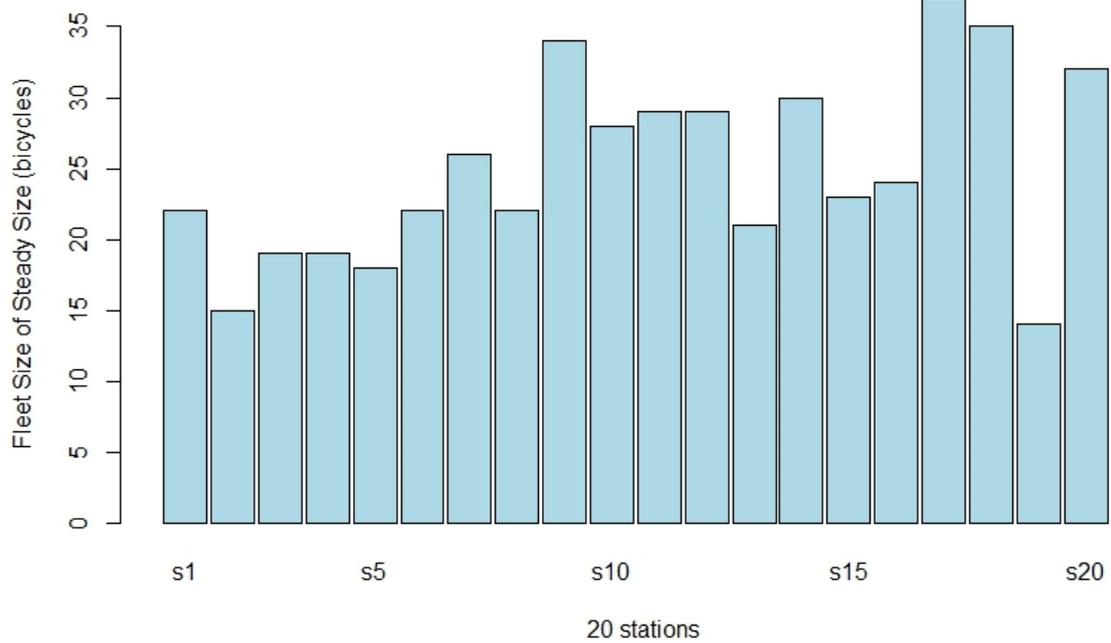


Figure 5. Statistical diagram of steady-state fleet size vector for dockless bike-sharing stations.

6. Compute the actual fleet size vector.

$$\gamma = (61 \ 2 \ 3 \ 27 \ 33 \ 14 \ 32 \ 51 \ 32 \ 2 \ 16 \ 45 \ 22 \ 6 \ 53 \ 41 \ 2 \ 21 \ 12 \ 25)$$

In Figure 6, the horizontal axis represents the stations and the vertical axis represents the actual fleet size of stations.

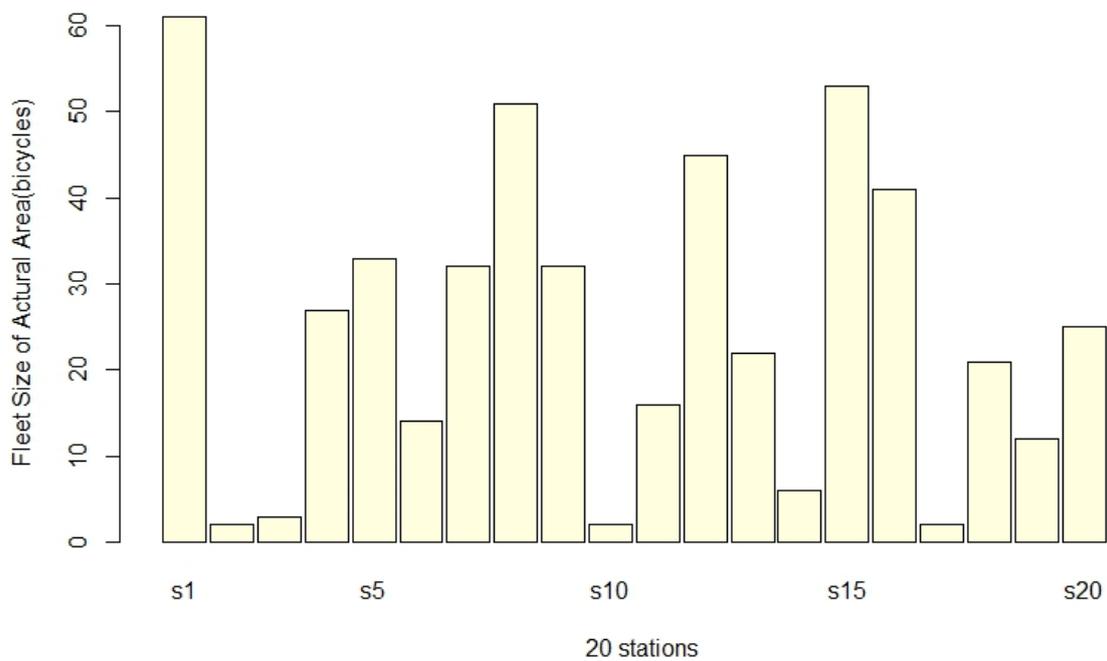


Figure 6. Statistical diagram of actual fleet size vector for dockless bike-sharing stations.

7. Compute the difference between the steady-state fleet size vector and the actual fleet size vector.

In Figure 7, the horizontal axis represents the stations and the vertical axis represents the difference between the steady-state fleet size vector and the actual fleet size vector.

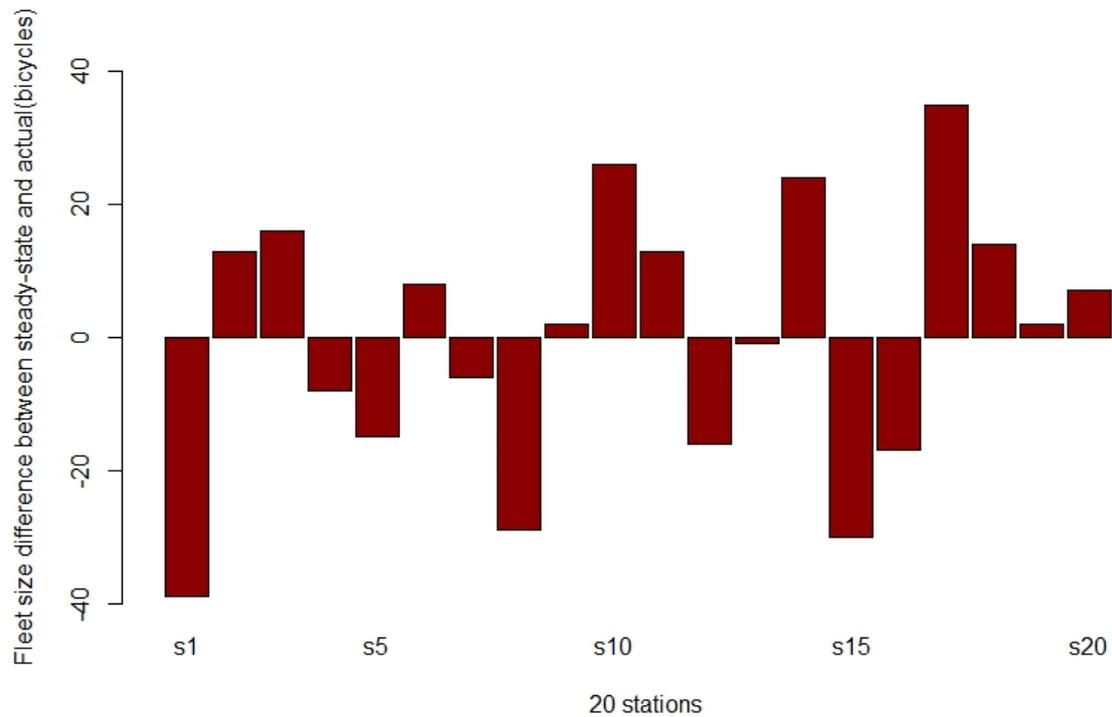


Figure 7. Statistical diagram of difference between the steady-state fleet size vector and the actual fleet size vector.

8. Construct the moving-out stations number vector and moving-in stations number vector.

(1) Moving-out stations sequence vector

$$\omega = (2\ 3\ 6\ 9\ 10\ 11\ 14\ 17\ 18\ 19\ 20)$$

(2) Moving-in stations sequence vector

$$v = (1\ 4\ 5\ 7\ 8\ 12\ 13\ 15\ 16)$$

9. Compute the distance matrix of the moving-out and the moving-in stations.

Taking the maximum value of the distance in the matrix as the denominator, the elements in the matrix are respectively used as molecules, and after dividing, the following distance matrix is constructed, which is showed in Figure 8

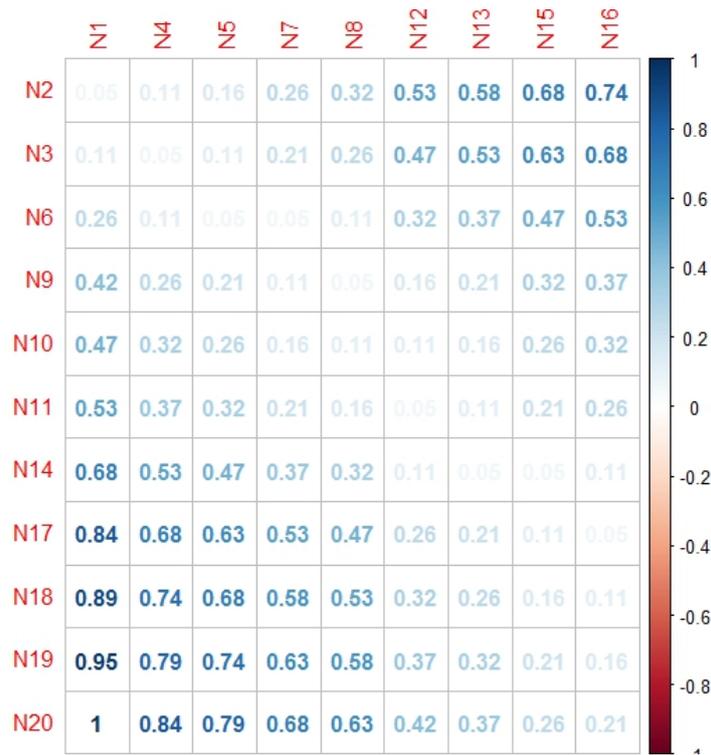


Figure 8. Statistical diagram of distance for moving-out and moving-in stations.

10. Solve the feasible solution of bicycle rebalancing.

In Figure 9, the number in the arc is the number of bicycle rebalancing, and the number in parentheses is the station serial number.

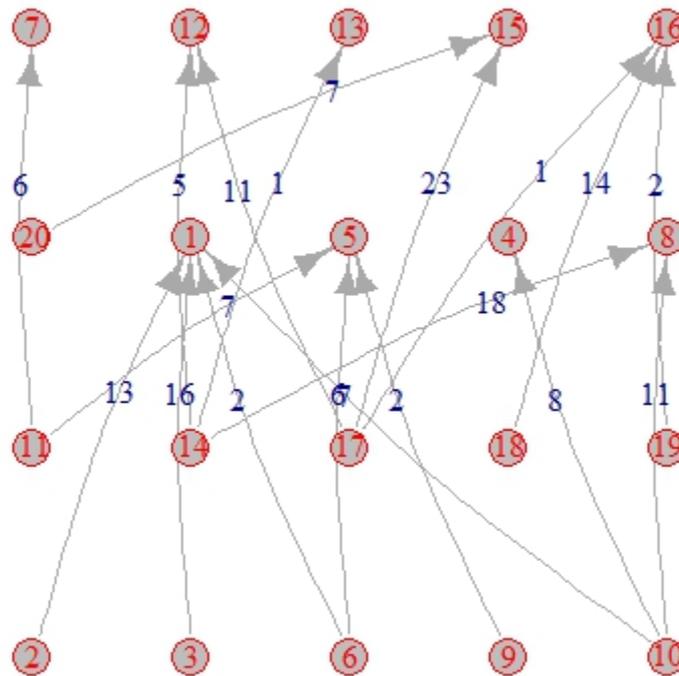


Figure 9. Rebalancing chart of the moving-out stations and moving-in stations.

In order to further verify the feasibility of the proposed algorithm, under the different total fleet size of shared bicycles and the number of bike-sharing stations, this paper computes and compares

the steady-state fleet size of the bike-sharing station and the convergence speed, respectively. In the simulation, the number of stations is set to 300, 500, 700, 900, and 1100, respectively. The total fleet size of bike-sharing is set to 1×10^5 , 2×10^5 , 3×10^5 and 4×10^5 , respectively. The adjustment parameter is taken as 0.96. The steady-state probability is computed 1000 times for a total of 20 simulation combinations for the bike-sharing station and the total fleet size. The average convergence speed is computed, see Table 3 for details.

Table 3. Simulation data steady-state probability solution iteration number statistics.

Number of Stations	Total Fleet Size			
	1×10^5	2×10^5	3×10^5	4×10^5
300	6.3	7	7	7
500	6	6	6	6.7
700	6	6	6	6
900	5.5	6	6	6
1100	5	6	6	6

Experiments show that when the total fleet size of shared bicycles in the city is constant, the more the number of stations, the fewer the number of iterations are required to reach the steady-state, that is, the less time is needed for the fleet size to reach the steady-state. Theoretically, this conclusion can also be proved. From Equation (7), when n is larger, the $0.04 \frac{1}{n} e^T (0.96^{k-1} \mathbf{P}'^{(k-1)} + 0.96^{k-2} \mathbf{P}'^{(k-2)} + \dots + 0.96 \mathbf{P}' + 1)$ in formula (7) is closer to zero, Equation (7) gets closer to Equation (8). Then, the less the factors affecting the steady-state of dockless bike-sharing, only affected by $\pi^{(0)} \mathbf{P}'^{(k)}$, so the number of iterations required to reach the steady-state is relatively reduced.

On the basis of solving the steady-state fleet size vector, this paper uses the Lingo software to analyze the linear programming of bicycle rebalancing in four scenarios with 300 stations and 1×10^5 , 2×10^5 , 3×10^5 , 4×10^5 bicycle numbers.

The choice of θ value is set based on the actual travel data of the dockless bike-sharing. In a general sense, if $\theta = 1$, it means that the shared bicycles are riding to the closest stations within 10 km, and the peripheral stations more than 10 km away are not used as the riding destination. This is an ideal situation, and the transition probability sparse matrix \mathbf{P}' is an irreducible matrix and does not need to be rank-one updating. If $\theta \rightarrow 0$, it means that the shared bicycles are almost all riding to the peripheral stations 10 km away, and the close stations within 10 km are not used as the riding destinations, which is obviously inconsistent with the facts.

The value of θ is chosen close to one. On the one hand, when $\theta \rightarrow 1$, the change in the probability value of the non-zero element in \mathbf{P}' is very small. It is actually consistent with the riding of the bike-sharing, and there are few users riding to the peripheral stations. On the other hand, the element updating control of the transition probability sparse matrix \mathbf{P}' is within an acceptable range. In this paper 3.1.4, after the rank-one updating, the probability value of the non-zero element P_{jk} in the sparse matrix \mathbf{P}' changes to $(1 - \theta) \left(\frac{1}{n} - P_{jk} \right)$. If the number of shared bicycle stations in the city is large (such as $n \geq 100$), then the absolute value of the above formula $|\Delta P| \approx (1 - \theta) P_{jk}$, taking $\theta = 0.96$ selected in this paper as an example, $|\Delta P| \approx 0.04 P_{jk}$, the probability value of non-zero elements changes very little.

If the peripheral stations are used more frequently, the difference between θ and one is larger, and the updating to \mathbf{P}' is also larger, thus the deviation between the rank-one updating method and the actual situation is more obvious. The experiment shows that the difference between θ and one is greater than 0.1, therefore this algorithm is not recommended.

5. Conclusions

In this paper, the traditional Markov chain steady-state probability vector solution method is improved, and the "Markov chain dockless bike-sharing fleet size solving method" is given. This

method combines the actual application of bicycles to construct a transition probability sparse matrix. For the sparse matrix may be reducibility, the method of arguing irreducibility based on graph theory and the general iterative method for the steady-state fleet size are given. For the reducibility sparse matrices, the rank-one updating method is used to conduct the random prime matrix to meet the requirements of steady-state fleet size solving, the iterative method of the solution is given and the convergence speed of the method is analyzed. On this basis, the construction method of the bicycle initial value and transition probability matrix in multiple time periods in one day is further analyzed, which makes the algorithm that can solve the station probability fleet size vector in different time periods within one day. Finally, the linear programming method is used to analyze the bicycle rebalancing, and the algorithm is verified by the experiment.

With the algorithm, one takes into account the computation accuracy and speed, and has strong operability. The steady-state fleet size is used as the baseline value to enable the relevant companies to dynamically optimize the bicycle deploy fleet size, and strengthen the standardized management.

Author Contributions: Data curation, Juan Du; Formal analysis, Jin Liu; Methodology, Yong Zhai; Writing—review & editing, Hao Wu.

Funding: This research received no external funding.

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

Appendix A

Proof

$(n-2)! \sum_{k=0}^{n-2} \frac{n-k-1}{k!}$ is convergence series, and $(n-2)! \sum_{k=0}^{n-2} \frac{n-k-1}{k!} \geq (n-2)! \sum_{k=0}^{n-2} \frac{\frac{n}{2}}{k!}$, $n \geq 3$.

Proof process:

$(n-2)! \sum_{k=0}^{n-2} \frac{n-k-1}{k!}$ is positive series, and $u_k = (n-2)! \frac{n-k-1}{k!}$, $k \in [0, n-2]$,

Then

$$\lim_{n \rightarrow \infty} \frac{u_{k+1}}{u_k} = \lim_{n \rightarrow \infty} \frac{(n-2)! \frac{n-k-2}{(k+1)!}}{(n-2)! \frac{n-k-1}{k!}} = \lim_{n \rightarrow \infty} \frac{n-k-2}{(k+1)(n-k-1)} < 1.$$

Therefore, it can be known that this series converges from the D'Alembert's discriminant.

Set $a_i = n-k-1$, $i = k+1$, we get

$$\sum_{k=0}^{n-2} \frac{n-k-1}{k!} = \frac{a_1}{0!} + \frac{a_2}{1!} + \dots + \frac{a_{n-1}}{(n-2)!} \quad (A1)$$

The following relationship exists in formula (1):

$$a_i + a_{n-i} = n$$

Then

$$a_i - \frac{n}{2} = \frac{n}{2} - a_{n-i} \quad (A2)$$

1. When n is an odd number, the number $n-1$ of the series term in Equation (1) is an even number.

Let $j \in [1, \frac{n-1}{2}]$, we have

$$\begin{cases} (j-1)! = (n-1-j)!, n=3 \\ (j-1)! < (n-1-j)!, n>3 \end{cases} \quad (A3)$$

By Equations (2) and (3), we get

$$\begin{cases} \frac{a_j - \frac{n}{2}}{(j-1)!} = \frac{\frac{n}{2} - a_{n-j}}{(n-1-j)!}, n = 3 \\ \frac{a_j - \frac{n}{2}}{(j-1)!} > \frac{\frac{n}{2} - a_{n-j}}{(n-1-j)!}, n > 3 \end{cases}$$

Then

$$\begin{cases} \frac{a_j - \frac{n}{2}}{(j-1)!} - \frac{\frac{n}{2} - a_{n-j}}{(n-1-j)!} = 0, n = 3 \\ \frac{a_j - \frac{n}{2}}{(j-1)!} - \frac{\frac{n}{2} - a_{n-j}}{(n-1-j)!} > 0, n > 3 \end{cases}$$

and

$$\sum_{k=0}^{n-2} \frac{n-k-1}{k!} - \sum_{k=0}^{n-2} \frac{\frac{n}{2}}{k!} = \frac{a_1 - \frac{n}{2}}{0!} + \frac{a_2 - \frac{n}{2}}{1!} + \dots + \frac{a_{n-1} - \frac{n}{2}}{(n-2)!} = \sum_{j=1}^{\frac{n-1}{2}} \left(\frac{a_j - \frac{n}{2}}{(j-1)!} - \frac{\frac{n}{2} - a_{n-j}}{(n-1-j)!} \right)$$

Then

$$\begin{cases} \sum_{k=0}^{n-2} \frac{n-k-1}{k!} - \sum_{k=0}^{n-2} \frac{\frac{n}{2}}{k!} = 0, n = 3 \\ \sum_{k=0}^{n-2} \frac{n-k-1}{k!} - \sum_{k=0}^{n-2} \frac{\frac{n}{2}}{k!} > 0, n > 3 \end{cases}$$

2. When n is an even number of not less than four, the number of series terms $(n-1)$ in the formula (1) is an odd number.

Let $j \in \left[1, \frac{n-1}{2}\right]$, we have

$$\begin{aligned} \sum_{k=0}^{n-2} \frac{n-k-1}{k!} - \sum_{k=0}^{n-2} \frac{\frac{n}{2}}{k!} &= \sum_{j=1, j \neq \frac{n}{2}}^{\frac{n-1}{2}} \left(\frac{a_j - \frac{n}{2}}{(j-1)!} - \frac{\frac{n}{2} - a_{n-j}}{(n-1-j)!} \right) + \frac{a_{\frac{n}{2}} - \frac{n}{2}}{\left(\frac{n}{2}-1\right)!} \\ &\because a_{\frac{n}{2}} = \frac{n}{2} \\ \therefore \frac{a_{\frac{n}{2}} - \frac{n}{2}}{\left(\frac{n}{2}-1\right)!} &= 0 \text{ and } \sum_{j=1, j \neq \frac{n}{2}}^{\frac{n-1}{2}} \left(\frac{a_j - \frac{n}{2}}{(j-1)!} - \frac{\frac{n}{2} - a_{n-j}}{(n-1-j)!} \right) \end{aligned}$$

Then

$$\sum_{k=0}^{n-2} \frac{n-k-1}{k!} - \sum_{k=0}^{n-2} \frac{\frac{n}{2}}{k!} > 0.$$

In summary

$$\sum_{k=0}^{n-2} \frac{n-k-1}{k!} \geq \sum_{k=0}^{n-2} \frac{\frac{n}{2}}{k!}, n \geq 3.$$

Then

$$(n-2)! \sum_{k=0}^{n-2} \frac{n-k-1}{k!} \geq (n-2)! \sum_{k=0}^{n-2} \frac{\frac{n}{2}}{k!}, n \geq 3.$$

References

1. The Bike-Sharing Blog. Available online: <http://bike-sharing.blogspot.com/> (accessed on 26 April 2019).
2. Shaheen, S.; Guzman, S.; Zhang, H. Bikesharing in Europe, the Americas, and Asia: Past, present, and future. *Transp. Res. Rec.* **2010**, *2143*, 159–167. [CrossRef]
3. Fishman, E. Bikeshare: A Review of Recent Literature. *Urban. Transport. China* **2016**, *36*, 92–113. [CrossRef]
4. Shaheen, S.A.; Martin, E.W.; Cohen, A.P.; Finson, R.S. *Public Bikesharing in North America: Early Operator and User Understanding*; Quadriga; Mineta Transportation Institute: San Jose, CA, USA, 2012.
5. Social Bicycles. Available online: <http://socialbicycles.com/> (accessed on 26 April 2019).
6. The Past and Present of Bike-Sharing—Incomplete Competing Analysis. Available online: <https://www.jianshu.com/p/68c0d1a3208e> (accessed on 26 April 2019).
7. Zhou, Y. The transportation system of Hangzhou’s public bicycle. *Traffic Transp.* **2011**, *27*, 36–37.
8. Li, J. Problems and suggestions for bike-sharing development. *Manag. Manag.* **2017**, *9*, 9–11.

9. China Bike-Sharing Industry Research Report. 2017–2021. Available online: <http://www.pday.com.cn/Htmls/Report/201704/24516157.html> (accessed on 26 April 2019).
10. What is Dockless Bike Share. Available online: <https://altaplanning.com/dockless-bike-share/> (accessed on 26 April 2019).
11. Banerjee, S.; Freund, D.; Lykouris, T. Multi-objective pricing for shared vehicle systems. *arXiv* **2016**, arXiv:1608.06819.
12. DeMaio, P. Bike-sharing: History, impacts, models of provision, and future. *J. Public Transp.* **2009**, *12*, 41–56. [[CrossRef](#)]
13. Chemla, D.; Meunier, F.; Wolfler Calvo, R. Bike sharing systems: Solving the static rebalancing problem. *Discret. Optim.* **2013**, *10*, 120–146. [[CrossRef](#)]
14. Faghieh-Imani, A.; Hampshire, R.; Marla, L.; Eluru, N. An empirical analysis of bike sharing usage and rebalancing: Evidence from barcelona and seville. *Transp. Res. Part. A Policy Pract.* **2017**, *97*, 177–191. [[CrossRef](#)]
15. Dell’Amico, M.; Hadjicostantinou, E.; Iori, M.; Novellani, S. The bike sharing rebalancing problem: Mathematical formulations and benchmark instances. *Omega* **2014**, *45*, 7–19. [[CrossRef](#)]
16. García-Palomares, J.C.; Gutiérrez, J.; Latorre, M. Optimizing the location of stations in bike-sharing programs: A GIS approach. *Appl. Geogr.* **2012**, *35*, 235–246. [[CrossRef](#)]
17. Martinez, L.M.; Caetano, L.; Eiró, T.; Cruz, F. An optimisation algorithm to establish the location of stations of a mixed fleet biking system: An application to the city of Lisbon. *Proc. Soc. Behav. Sci.* **2012**, *54*, 513–524. [[CrossRef](#)]
18. Raviv, T.; Tzur, M.; Forma, I.A. Static repositioning in a bike-sharing system: Models and solution approaches. *EURO J. Transp. Logist.* **2013**, *2*, 187–229. [[CrossRef](#)]
19. Rainer-Harbach, M.; Papazek, P.; Hu, B.; Raidl, G.R. *Balancing Bicycle Sharing Systems: A Variable Neighborhood Search Approach*; Lecture Notes in Computer Science; Springer: Heidelberg, Germany, 2013; pp. 121–132.
20. Kloimüller, C.; Papazek, P.; Hu, B.; Raidl, G.R. Balancing Bicycle Sharing Systems: An Approach for the Dynamic Case. In *European Conference on Evolutionary Computation in Combinatorial Optimization*; Springer: Heidelberg, Germany, 2014; pp. 73–84.
21. McKenzie, G. Docked vs. Dockless Bike-sharing: Contrasting Spatiotemporal Patterns (Short Paper). In Proceedings of the 10th International Conference on Geographic Information Science (GIScience 2018), Melbourne, VIC, Australia, 28–31 August 2018; Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik: Saarland, Deutschland; pp. 336–348.
22. Mobike About. Available online: <https://mobike.com/global/about> (accessed on 26 April 2019).
23. Xiong, S. An Internet-Based Short-Distance Travel Solution in China. Bachelor’s Thesis in Business, Lahti University of Applied Sciences, Lahti, Finland, 2017.
24. Luo, Q.; Dou, X.; Di, X.; Hampshire, R.C. Multimodal Connections between Dockless Bikesharing and Ride-Hailing: An Empirical Study in New York City. In Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 4–7 November 2018; IEEE: Piscataway, NJ, USA, 2018.
25. Cui, W. The Effects of Urban Density on the Efficiency of Dockless Bike Sharing System—A Case Study of Beijing, China. Master’s Theses, Arizona State University, Tempe, AZ, USA, 2018.
26. Pan, L.; Cai, Q.; Fang, Z.; Tang, P.; Huang, L. A Deep Reinforcement Learning Framework for Rebalancing Dockless Bike Sharing Systems. *arXiv* **2018**, arXiv:1802.04592.
27. Wang, G. The traffic value and development path of bike-sharing. *City* **2017**, *205*, 72–75.
28. Jiao, G. A brief analysis of the sustainable profit model of the Bike-sharing. *Consum. Guide* **2017**, *39*, 37–38.
29. Hu, J. User Behaviour Analysis and Optimization in Bike-Sharing Systems. China Master’s Theses, Zhejiang University, Hangzhou, China, 2018.
30. Yang, Y. Design and research on Crime Prevention Through Environmental about Bike-sharing. *Chin. Criminol. Rev.* **2017**, *9*, 67–74.
31. Zhai, Y.; Liu, J.; Du, J.; Chen, J. Solution to fleet size of dockless bike-sharing station based on matrix analysis. In Proceedings of the 2018 ISPRS TC IV Mid-term Symposium “3D Spatial Information Science—The Engine of Change”, Delft, The Netherlands, 1–5 October 2018.
32. Liu, C. *Stochastic Process*, 4th ed.; Huazhong University of Science and Technology Press: Wuhan, China, 2008; pp. 42–70.

33. Ching, W.K.; Ng, M.K. *Markov Chains. Models, Algorithms and Applications*; Springer: Boston, MA, USA, 2006; pp. 1–46.
34. Langville, A.N.; Meyer, C.D. *Google's PageRank and Beyond: The Science of Search Engine Rankings*; Princeton University Press: Princeton, NJ, USA, 2011; pp. 40–41, 172–174.
35. Langville, A.N.; Meyer, C.D. Updating Markov chains with an eye on Google's PageRank. *Siam J. Matrix Anal. Appl.* **2006**, *27*, 968–987. [[CrossRef](#)]
36. The Secret of Codoon's Bike-Sharing Data. Available online: https://www.codoon.com/iec/article/PwdL_gm_O5E= (accessed on 26 April 2019).
37. Tarjan, R. Depth-first search and linear graph algorithms. *SIAM J. Comput.* **1972**, *1*, 146–160. [[CrossRef](#)]
38. Stewart, W.J. *Introduction to the Numerical Solution of Markov Chains*; Princeton University Press: Princeton, NJ, USA, 1994; pp. 121–125.
39. Mao, D.; Hao, Z.; Wang, Y.; Fu, S. A Novel Dynamic Dispatching Method for Bicycle-Sharing System. *ISPRS Int. J. GeoInf.* **2019**, *8*, 117. [[CrossRef](#)]
40. Dantzig, G. *Linear Programming and Extensions*; Princeton University Press: Princeton, NJ, USA, 2016; pp. 35–36.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).