SOLUTION TO FLEET SIZE OF DOCKLESS BIKE-SHARING STATION BASED ON MATRIX ANALYSIS

Yong Zhai [1], Jin Liu [1], Juan Du [1], Jie Chen [1]

1 National Geomatics Center of China, 100830 Beijing, China-(zhaiyong,liujin,dujuan,chenjie)@ngcc.cn

Commission IV, WG IV/4

KEY WORDS: Markov chain, steady-state fleet size, transition probability, random matrix, power method, rank-one updating method

ABSTRACT:

Aiming at the problems of the lack of reasonable judgment of fleet size and non-optimization of rebalancing for dockless bike-sharing station, based on the usage characteristics of dockless bike-sharing, this paper demonstrates that the Markov chain is suitable for the analysis of the fleet size of station. It is concluded that dockless bike-sharing Markov chain probability limit state (steady-state) only exists and is independent of the initial probability distribution. On that basis, this paper analyses the difficulty of the transition probability matrix parameter statistics and the power method of the bike-sharing Markov chain, and constructs the transition probability sparse matrix in order to reduce computational complexity. Since the sparse matrices may be reducible, the rank-one updating method is used to construct the transition probability random prime matrix to meet the requirements of steady-state size calculation. An iterative method for solving the steady-state probability is therefore given and the convergence speed of the method is analysed. In order to improve the practicability of the algorithm, the paper further analyses the construction methods of the initial values of the dockless bike-sharing and the transition probability matrices at different time periods in a day. Finally, the algorithm is verified with practical and simulation data. The results of the algorithm can be used as a baseline reference data to dynamically optimize the fleet size of dockless bike-sharing station operated by bike-sharing companies for strengthening standardized management.

1. INTRODUCTION

Since 2016, bike-sharing (MetroBike, 2018) have been flourished in China at an impressive speed (Shao et al., 2017). As we known, the Mobike company, which has the world's largest mobile IoT network with over nine million shared bicycles in operation serving up to 30 million rides daily (Mobike, 2018a). Shared bicycles deployed into market by bike-sharing companies since 2016 are mainly dockless type (wikipedia, 2018), allowing users to pick up freely (referred to as renting) or return a bicycle in public spaces of the city.

In the research of this paper, the public spaces of dockless bike-sharing placement are defined as stations. For the convenience of calculation, a city can be divided into many open square spaces with equal area, and each space can be regarded as a station of dockless bike-sharing. The fleet size of bike-sharing station mentioned below refers to the number of bicycles in the station. The total fleet size of the dockless bike-sharing is the total number of shared bicycles put into a city.

The emergence of dockless bike-sharing is an innovative mode of sharing economy in the Internet era, which improves "the first mile/last mile" connection to other modes of transit, and lessen the environmental impacts of our transport activities (DeMaio, 2009). However, on the other hand, the disorder of dockless bike-sharing put into market has become increasingly prominent, some bike-sharing companies to win the competition by quantity, and expansion presents chaotic. The problems of fleet size and rebalancing of the dockless bike-sharing stations are mainly as follows:

(1) There are too many bicycles in some stations, the problems of idleness are prominent, which even can block pedestrian and cars normal traffic. At the same time, the number of bicycles in some other stations is too small to meet the needs of users.

(2) The bike-sharing replenishment or repositioning transport response is not timely, and the rationality of the rebalancing is debatable.

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 in order to achieve the goal of standardized management.

At present, in China and other countries there are many related researches on the fleet size and rebalancing of docked bike-sharing. However, due to the operation time is relatively short, there are few research literatures on the fleet size and rebalancing of dockless bike-sharing operated by companies. Researchers pay more attention to the traffic value and development path of dockless bike-sharing (Wang, 2017), the profitability approach (Jiao, 2017), behavioural norms (Hu, 2018) and the design of crime prevention environment (Yang, 2017). There are significantly difference between dockless and docked bike-sharing in station location, renting and returning mode, operation and charging mode, and so on. The dockless bike-sharing has following characteristics: flexibility, convenience and randomness, larger user group, higher frequency of use, wider coverage. It is more difficult to study the fleet size and rebalancing of dockless bike-sharing stations, and the methods is different.

Assuming the location of all stations has been determined, the paper uses matrix analysis to solve the optimal value of the fleet size of dockless bike-sharing station. In order to eliminate the interference of major economic activities, bad weather, and user

personal preferences for different dockless bike-sharing brand, this paper mainly analyses the fleet size and rebalancing of dockless bike-sharing stations from the same company under daily use.

2. MARKOV CHAIN APPLICABILITY ANALYSIS

2.1 Properties of bike-sharing Markov Chain

Bike-sharing provides people with a convenient way for commuting by shared bicycles among users, and solves the “first/last mile” 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. It is also more common for users to rent bicycles between stations that are far away from each other. In this paper, theoretically assumed that there is connectivity or bicycle state transition relationships between any two stations.

Assuming that a user rents the shared bicycle from the station j within the time period $t_0 - t_1$ and returns the rented bicycle at the station k. The state of the station k during the transfer is only related to the station j, regardless of the state of other stations in the same interval, and independent of the state of this station and other stations before the time. From this basic sense, the state relationship of shared bicycle stations is in accordance with the basic conditions of Markov random process application. The main characteristic of the dynamic feature of Markov random processes is that the probability distribution of future states depends only on the current state, and has nothing to do with the process of reaching the current state (Liu, 2008). For the problem of fleet size and rebalancing of dockless bike-sharing stations, the paper mainly studies the time-discrete Markov random process.

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

(1) The user at the station j rents a shared bicycle ride to any other stations that exists n-1 kind of possible (mutually exclusive), each of which has two states (e.g, the shared bicycle at station j and station k). Set the shared bicycle moves 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}, $P_{jk} = P\{X_k \mid X_j\}$, $k \neq j$.

(2) The shared bicycle at the station j is not rented by the user or returned to the station j after the user used, 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 shown in Figure 1.

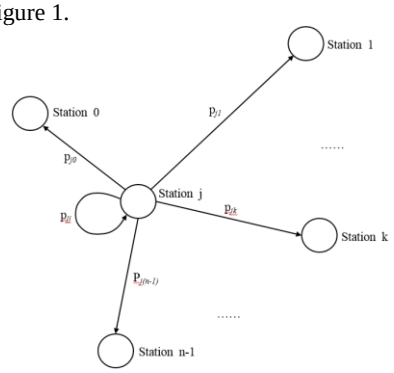

Figure 1. Transition states of station j

For the probability of dockless shared bicycle 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. P_{jj} can be regarded as the ratio of the average number of bicycles that not rented or returned by users in the multiple time periods to the average number of bicycles in station j.

In summary, the probability P_{jk} that the user rides from station j to other stations and the probability P_{jj} that is still returned to station j after rented or no user rented are accumulated to 1. $\sum_{k=0}^{n-1} P_{jk} = P_{j0} + P_{j1} + \cdots P_{jj} + \cdots P_{j(n-1)} = 1, P_{jk} > 0$ (To ensure connectivity between stations, $P_{jk} \neq 0$), $k = 0,1,2,\cdots,n-1$.

By analogy, the transition state of n stations is shown in Figure 2.

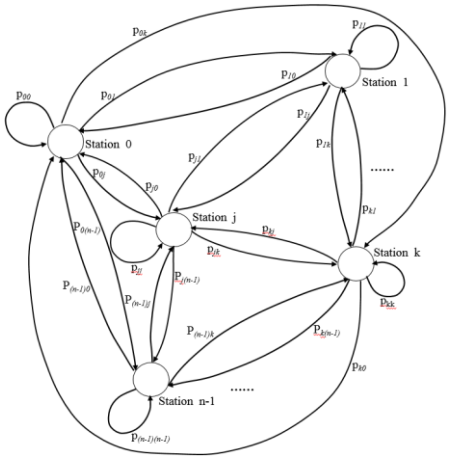

Figure 2. Transition states of all stations

The transition probability matrix of all n stations is constructed as

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

$$\begin{cases} P_{jj} = 1 - \sum_{k=1}^{n-1} P_{0k}, & when\ j = 0 \\ P_{jj} = 1 - \sum_{k=0}^{n-2} P_{(n-1)k}, & when\ j = n-1 \\ P_{jj} = 1 - \sum_{k=0}^{j-1} P_{jk} - \sum_{k=j+1}^{n-1} P_{jk}, & when\ j \neq 0, n-1 \end{cases}$$

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

At a certain time, the ratio of the number of bicycles owned by each bicycle station to the total number 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 at time t_1 is $\pi^{(1)}$.

From the nature of the Markov chain, we can obtained $\pi^{(1)}=\pi^{(0)}P$.

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

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

where P = transition probability matrix
$\pi^{(k)}$ = probability vector after k steps transfer
$\pi^{(k-1)}$ = probability vector after k-1 steps transfer

Iterate over equation (1) and get

$$\pi^{(k)} = \pi^{(k-1)}p=\pi^{(k-2)}p^2=\cdots=\pi^{(0)}p^k \tag{2}$$

2.2 Proof of the existence of steady state of bike-sharing Markov Chain

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

(1) Irreducible. Let B be a non-empty subset of the state space I . If $i \in B$ and $k \notin B$, then connectivity probability of j and k $P_{ik}=0$, B is called a closed set. If all the states in B are interconnected, B is called an irreducible closed set (Liu, 2008). If Markov chain's state space is an irreducible closed set, it is called irreducible (Liu, 2008). 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. The bike-sharing Markov chain theoretically has irreducible properties (Ching et al., 2006).

(2) 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. According to Carl D. Meyer's research, if random matrix is irreducible and has at least one positive diagonal element, then it is a prime matrix (Langville et al., 2011). Investigate the transition probability matrix P , which is an irreducible non-negative square matrix with n diagonal element (P_{jj}) . Investigation of the use of dockless bike-sharing, all the bicycles of 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, only at this time all the diagonal elements of P are 0, which is inconsistent from the practical situation, non-zero diagonal elements are more prevalent. Therefore, the dockless bike-sharing Markov chain is aperiodic (non-periodic).

(3) 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 recurrence. When the average return time of state j is finite, it is called positive-recurrence. Liu (2008) has shown that irreducible finite Markov chains must be recurrence in random process. It has been discussed above that the bike-sharing Markov chain has the finite and irreducible properties, so all the states in the bike-sharing Markov chain are 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: 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 is big enough, we have

$$\lim_{k\to\infty}\pi^{(k)} = \pi \tag{3}$$

where $\pi^{(k)}$ = the probability vector after k steps transfer
π = a steady state probability vector

From formula (2), we have

$$\lim_{k\to\infty}\pi^{(k)} = \lim_{k\to\infty}(\pi^{(0)}P^k) = \pi \tag{4}$$

Therefore, the main idea of the algorithm is based on multiplying the initial probability vector of the dockless bike-sharing station by the higher power of the transition probability matrix, that is $\pi^{(0)}P^k$. The calculated 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 multiplied by the total number of shared bicycles respectively, thereby obtaining 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 power method.

3. DIFFICULTY ANALYSIS OF BIKE-SHARING STEADY-STATE SOLUTION

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 the stations in the city during the time period t_0-t_1 . For any station i , it is necessary to count the specific destination station for each user riding out of the bicycle. And for any station j , 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, and the workload of observation and calculation increases exponentially with the increase of stations. If the number of dockless bike-sharing stations in first-tier cities is thousands, the observation statistics of the transfer probability value is greater than one million $(\geq 1000^2)$. For those bike-sharing companies that cannot monitor, observe and count the bike-sharing data online, the workload is huge and the operability is poor. If the number of bike-sharing stations in tier 2 cities is hundreds, the observation statistics of the transfer probability value is tens of thousands to hundreds of thousands, and the workload is considerable. In addition, it is also necessary to calculate the initial value of the fleet size of each station in the same segment, so n stations have n numbers in total. And you 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 .The statistical workload is very huge, which is the

main bottleneck for solving the steady-state probability vector of dockless bike-sharing stations.

Solving equation (4) by power method usually requires $O(n^2)$ sub-vector-matrix multiplication calculation (n is the total number of bike-sharing stations), the amount of calculation is large. Markov chain steady-state probability vector solving is usually complicated (Liu, 2012).

In summary, combined with the fact that the accuracy of the requirements is not critical but the reaction speed is high, the dockless bike-sharing operation needs to simplify and optimize the algorithm, reduce the statistical workload, and improve the solution speed to improve feasibility.

4. FLEET SIZE ANALYSIS OF DOCKLESS BIKE-SHARING STATION

4.1 Sparse matrix construction

Based on the practical observation of the usage of dockless bike-sharing, it is found that for any station j, the riding of shared bicycle is mainly performed between 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 shown in Table 1(GuDong, 2017).

Serial number	Single ride distance (d)	The proportion of users
1	d≤3km	41.6%
2	3<d≤5km	39.4%
3	5<d≤10km	15.2%
4	d>10km	3.8%
Sum		100%

Table 1. Bike-sharing user single riding distance statistics

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

(1) Close station. For any station j, the station that is within 10 km away from it is a close station.

(2) Peripheral station. For any station j, the station that is more than 10 km away from it is a peripheral station. In this paper, the ratio of the number of peripheral stations to the total number of stations is 0.04 on average.

(3) Unrelated station. In this paper, the stations with no bicycle transfer relationship are called unrelated stations, and the transition probability is set to 0.

Based on above analysis, a new dockless bike-sharing transition probability matrix P' is constructed, which is a sparse matrix. 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, $\hat{m} \ll n$, $\hat{m} \times n \ll n^2$. It means that the computational complexity of vector-matrix multiplication will tend to be $O(n)$. Therefore, the vector-matrix multiplication based on P' is far less than $O(n^2)$.

This method can be referred to as "sparse matrix construction solution method". Using this method, 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 and 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 $\hat{m} \times n + n + 1$, which is significantly smaller than the traditional total number of statistics $n^2 + n + 1$.

4.2 Steady-state solution

Since the station interoperability is affected due to the unreachability of some stations, the transition probability matrix should be sparse. The irreducible nature of the transition matrix may be changed to reducible. The steady-state condition of the bike-sharing is destroyed, and the steady-state probability maynot be solved. At this point, the rank-one updating method needs to be performed on the sparse matrix to satisfy the irreducible condition.

From the perspective of dockless bike-sharing applications, 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 above 10 km (not exceed 4%) is evenly distributed to all stations. Under the premise of the sum of the elements of each row of the matrix is one, a new random matrix that satisfy the steady-state probability of dockless bike-sharing is constructed. To simplify the calculation and minimize the effect on the state transition relationship in the transition probability sparse matrix P' , P' can be multiplied by one adjustment parameter θ (that is $\theta P'$, $0<\theta<1$, and θ close to 1) to decrease the non-zero elements appropriately. Then the sum of the elements of each row in the matrix P' is $1\times\theta=\theta$. Next, each element in P' add $(1-\theta)/n$ separately, and the sum of the increments in each row is $\left[(1-\theta)/n\right]\times n = 1-\theta$, 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

$$C=\theta P'+\frac{1-\theta}{n}ee^T \tag{5}$$

e^T is the transpose for the unit column vector e . The rank of the matrix formed by $\left[(1-\theta)/n\right]ee^T$ is 1, 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 take 0.96, and the calculation result of the rank-one updating part is $n\times n$ square matrix. We have

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

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

After the rank-one updating, the probability value of the non-zero element P'_{jk} in the sparse matrix P' changes as follows

$$\begin{aligned}\Delta P' &= (P'_{jk}\theta + (1-\theta)\frac{1}{n}) - P'_{jk} \\ &= (0.96P'_{jk} + 0.04\frac{1}{n}) - P'_{jk} \\ &= 0.04(\frac{1}{n} - P'_{jk})\end{aligned}$$

If the number of bike-sharing stations in a city is large (e.g. $n \geq 100$), then the absolute value of the above formula is $|\Delta P'| \approx 0.04P'_{jk}$, and the probability value of non-zero elements changes minimally.

In addition, after the rank-one updating, the probability value of each zero element is changed to $(1-\theta)/n = 0.04/n$. If the number of bike-sharing stations in the city is $n \geq 100$, the corrected probability value of each zero element is not more than 0.0004.

The larger the number of bike-sharing stations, the smaller the impact of the rank-one updating on P' elements. Therefore, after the rank-one updating of the sparse matrix P' , the constructed new transition probability matrix C does not change the main state transition relationship (riding relationship) between stations. For a bike-sharing operation with a low accuracy requirement and a high response speed, the solution can be reasonable.

When calculating the steady-state probability of C by power method, we have

$$\begin{aligned}&\pi^{(k)} \\ &= \pi^{(k-1)}C \\ &= \pi^{(k-1)}(\theta P' + (1-\theta)\frac{1}{n}ee^T) \\ &= \theta\pi^{(k-1)}P' + (1-\theta)\frac{1}{n}\pi^{(k-1)}ee^T \qquad (6) \\ &\because \pi^{(k-1)}e = 1 \\ &\therefore \pi^{(k)} = \theta\pi^{(k-1)}P' + (1-\theta)\frac{1}{n}e^T\end{aligned}$$

Although the matrix C is a dense matrix, however, it can be known from equation (6) that we actually perform vector-matrix operations on the original sparse matrix P' when performing 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 calculation, the time complexity of this algorithm is still $O(n)$.

Continue iterating on the $\pi^{(k)}$ to get

$$\begin{aligned}&\pi^{(k)} \\ &= \theta\pi^{(k-1)}P' + (1-\theta)\frac{1}{n}e^T \\ &= \theta(\theta\pi^{(k-2)}P' + (1-\theta)\frac{1}{n}e^T)\ P' + (1-\theta)\frac{1}{n}e^T \\ &= \theta^2\pi^{(k-2)}P'^{(2)} + \theta(1-\theta)\frac{1}{n}e^TP' + (1-\theta)\frac{1}{n}e^T \\ &\vdots \\ &= \theta^k\pi^{(0)}P'^{(k)} + \theta^{k-1}(1-\theta)\frac{1}{n}e^TP'^{(k-1)} + \theta^{k-2}(1-\theta)\frac{1}{n}e^TP'^{(k-2)} \\ &\quad + \cdots + \theta(1-\theta)\frac{1}{n}e^TP' + (1-\theta)\frac{1}{n}e^T \\ &= \theta^k\pi^{(0)}P'^{(k)} + (1-\theta)\frac{1}{n}e^T(\theta^{k-1}P'^{(k-1)} + \theta^{k-2}P'^{(k-2)} + \cdots + \theta P' + 1)\end{aligned}$$

Substituting $\theta = 0.96$ into the above formula,

$$\begin{aligned}\pi^{(k)} &= 0.96^k\pi^{(0)}P'^{(k)} \\ &\quad + 0.04\frac{1}{n}e^T(0.96^{k-1}P'^{(k-1)} + 0.96^{k-2}P'^{(k-2)} + \cdots + 0.96P' + 1)\end{aligned} \qquad (7)$$

When n is large enough, we have

$$\pi^{(k)} \approx 0.96^k\pi^{(0)}P'^{(k)} \qquad (8)$$

4.3 Convergence determination method

The dockless bike-sharing transition probability matrix P' is a random matrix, its matrix norm $\|P'\|_\infty = \max_{1\leq i\leq n}\sum_{j=1}^{n}|P'_{ij}| = 1$. 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 any of its matrix norms, that is $\rho(P') \leq \|P'\|_\infty = 1$. It is also defined that the spectral radius $\rho(P')$ is the largest value of the eigenvalues of the matrix P' (the absolute value of the dominant eigenvalue), then it is known that the absolute value of the P' dominant eigenvalue is $|\lambda_1| \leq 1$.

At the same time, since P' is a random matrix, $Pe = e = 1\times e$, that is, 1 is a eigenvalue of P' .

From the above, the dominant eigenvalue of the bike-sharing transition probability matrix P' is $\lambda_1 = 1$. For a random prime matrix, the dominant eigenvalue is unique (Stewart et al., 1994), therefore, without loss of generality, we can make the following n-1 eigenvalues have the following relationship $|\lambda_2| \geq |\lambda_3| \geq \ldots \geq |\lambda_n|$ .Then $1 > |\lambda_2| \geq |\lambda_3| \geq \ldots \geq |\lambda_n|$.

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

Based on the random prime matrix C , we calculate the steady-state probability vector of the bike-sharing by power method, and its convergence rate is $|\mu_2|^k \to 0$ (Langville et al., 2011). Since $|\mu_2| \leq \theta < 1$, the convergence rate does not exceed $|\theta|^k \to 0$. If the precision of the $\pi^{(k)}$ calculation 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.

Combining the iterative process of power method, when calculating actually the steady-state probability vector, the steady-state can be judged by calculating the absolute value of the difference between the corresponding elements of adjacent probability vectors. Set

$$\pi^{(k)} = (x_1, x_2, \cdots, x_n)$$

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

When every absolute value of the difference between the corresponding elements of the two adjacent vectors are all less than a sufficiently small constant α , we have

$$(|y_1\text{-}x_1| \le \alpha) \cup (|y_2\text{-}x_2| \le \alpha) \cup \cdots \cup (|y_n\text{-}x_n| \le \alpha) \ , \alpha = \frac{1}{S}$$

where S = the total fleet size of dockless shared bicycles put into a city

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

5. ALGORITHM IMPLEMENTATION

In practical calculations, $\pi^{(0)}$ is as the initial probability vector, it can take the value $(1/n,\ 1/n,\ \cdots 1/n)$. After solving the steady-state probability vector π by the power method, the total number of bicycles S and π are multiplied to obtain the steady-state fleet size vector of dockless bike-sharing stations. If the minimum time period of the bike-sharing state transition is in units of days, 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.

This section uses commercial data and simulation data to verify the feasibility and practical significance of the algorithm.

The commercial data is provided by Mobike company (Mobike, 2018b), which is more than 1.31 million lines. After the commercial data verification, the technical experts of Mobike company believe that the steady-state fleet size of the bike-sharing solved by this algorithm can be used as the baseline reference data in practical operation.

The simulation data is generated randomly by R language, and the matrix calculation is performed according to the algorithm of this paper to illustrate the feasibility of obtaining the steady-state fleet size of the station for decision support.

Just demonstrate the calculation process of the algorithm, the simulation experiment takes 5 stations as an example. Assuming the total size of bike-sharing is 500. The steady-state fleet size vector solution process is as follows

1) Building a total transition probability matrix of dockless bike-sharing.

$$P' = \begin{pmatrix} 0 & 0.4205 & 0 & 0 & 0.5795 \\ 0.2654 & 0 & 0 & 0.3075 & 0.4271 \\ 0 & 0 & 0 & 0.5767 & 0.4233 \\ 0.5280 & 0.0381 & 0.4339 & 0 & 0 \\ 0.8656 & 0 & 0.1344 & 0 & 0 \end{pmatrix}$$

2) Rank-one updating for P' , $C = \theta P' + [(1\text{-}\theta)/n] ee^T$ .The adjustment parameter θ is taken as 0.96. The results are as follows

$$C = \begin{pmatrix} 0.008 & 0.4117 & 0.008 & 0.008 & 0.5643 \\ 0.2628 & 0.008 & 0.008 & 0.3032 & 0.418 \\ 0.008 & 0.008 & 0.008 & 0.5617 & 0.4143 \\ 0.5148 & 0.0446 & 0.4245 & 0.008 & 0.008 \\ 0.839 & 0.008 & 0.1370 & 0.008 & 0.008 \end{pmatrix}$$

3) Calculating the steady-state probability vector of bike-sharing stations. On the basis of the rank-one updating, the steady-state probability vector is obtained by continuous iterations (the total number of iterations is equal to 18) as follows

$$\pi = (0.3508 \quad 0.1534 \quad 0.09 \quad 0.1031 \quad 0.3027)$$

4) Calculating the fleet size vector of dockless bike-sharing stations under the steady-state conditions. The steady-state fleet size vector of the bike-sharing is equal to the total size of the bicycles multiplied by the steady-state probability vector.

$$S * \pi = (175 \quad 77 \quad 45 \quad 52 \quad 151)$$

In order to further verify the feasibility of the proposed algorithm, under the different total size of shared bicycles and the number of bike-sharing stations, this paper calculates and compares the steady-state fleet size of bike-sharing station and the convergence time, 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 θ is taken as 0.96. The steady-state probability is calculated 1000 times for a total of 20 simulation scenarios for the bike-sharing station and the total fleet size. The steady-state fleet size vector is not displayed because its dimension is too large. The average convergence time is calculated, see Table 2 for details.

Transition probability matrix order	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

Table 2. The average convergence time of simulation data steady-state probability solution

When the number of bicycle stations in the city is constant, the experimental results show that the convergence time increases with the increase of the total fleet size of the bike-sharing. That means, when the number of stations is constant, the larger the total size of bike-sharing, the longer to the steady-state. When the total fleet size of bike-sharing is constant, the more number of stations, the less convergence time. That means, when the total size of bike-sharing is constant, the more the number of stations in the city, the shorter to the steady-state.

The simulation proves that the proposed algorithm is effective and can provide strategic support for bicycle operation.

In practical applications, the decision of fleet size for each bike-sharing station can refer to the calculation results under steady-state conditions, and resolve the rebalancing by considering factors such as station area.

If the calculated steady-state fleet size of bike-sharing station is greater than the maximum allowable number of bicycles in the area, then set j be such station.

(1) If P_{jj} is too large (e.g., $p_{jj} \geq 5\%$), which indicates that station j has more idle bicycles, so the idle bicycles need to be removed.

(2) If P_{jj} is within the normal range of values (e.g., $p_{jj} < 5\%$), which indicates that the bike-sharing steady-state fleet size of such bike-sharing stations exceeds the bearing capacity of the station area, and the number of bicycles in the station needs to be dynamically optimized. There are two cases:

1) The station has the characteristics that the bike-sharing number of riding in is greater than the number of riding out (For example, the station near the subway station in the residential area at the morning peak). With the number of bicycles gradually increase, the bike-sharing company should remove the bicycles in time to prevent the occupation of space for pedestrians or vehicles.

2) The station has the characteristics that the bike-sharing number of riding in is less than the number of riding out (For example, the station near the subway station in the office area at the morning peak). With the number of bicycles gradually decrease, the bike-sharing company should continue to supplement the number of bicycles.

It should be emphasized that if the shared bicycles are generally idle, it means that the total number of bicycles in the city exceeds the actual demand, which requires overall compression. It is advisable to control the average value of P_{jj} within 5% by removing the idle bicycle of the station.

This paper does not further analyse the problem of shared bicycle rebalancing, and only makes a few suggestions. In order to reduce the cost of the rebalancing, we should pay attention to the management of user behaviours. For users who park bicycles in disorder, they would be discouraged online. If the user does not comply with the dissuasion, online penalties can be imposed, such as increasing the rental fee to reduce the disorder; Users who actively use the disorderly parked bicycles, and pick up the bicycles from the congested stations or return them to the bicycles-deficient stations would be rewarded online. On this basis, the shared bicycle company uses static or dynamic bicycle repositioning problem algorithms to achieve reasonable rebalancing of bicycles.

6. ALGORITHM IMPROVEMENT

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, stable use at noon, and less used at night. Therefore, the working day time is divided as shown in Table 3.

State cycle	Time period
T_1	$[7:00-9:00)$
T_2	$[9:00\text{-}11:00)$
T_3	$[11:00-13:00)$
T_4	$[13:00\text{-}17:00)$
T_5	$[17:00-19:00)$
T_6	$[19:00-22:00)$
T_7	$[22:00$ to the next morning $7:00)$

Table 3. Bicycle Station State Change Cycle (24 hours)

The 24 hours in Table 3 is a cycle of changes in the state of the bike-sharing stations. For bike-sharing companies, the change of the fleet size of bike-sharing station in different time periods can provide an important basis for the formulation of the rebalancing strategy. Therefore, it is necessary to further improve the sparse matrix solution algorithm.

According to the analysis of Section 2, we can see that the change of bicycle state in the six time periods all meet the basic requirements of discrete-time Markov chain applications. Therefore, the algorithm can be further refined to dynamically calculate the transition probability matrix of each major time period of the day, and then calculate 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+1th day is $S_{T_i}^{(k+1)}, i \in (1,7)$. In the i-th time period, the number of rides out of the station j is $\mathrm{G}_{T_i}^{(k+1)}$ and the number of rides in is $I_{T_i}^{(k+1)}$. Then

$$S_{T_i}^{(k+1)} = S_{T_{(i-1)}}^{(k+1)} + I_{T_i}^{(k+1)} - \mathrm{G}_{T_i}^{(k+1)}, i \neq 1$$

$$S_{T_i}^{(k+1)} = S_{T_6}^{(k)} + I_{T_7}^{(k)} - \mathrm{G}_{T_7}^{(k)}, i = 1$$

Known the initial value S_{T_i} of bike-sharing for a certain time period in the station j, and the number of shared bicycles G_{T_i} from the station j to the station k, then the bike-sharing transition probability of the station j to station k in this time period is $P_{T_i-jk} = \mathrm{G}_{T_i} / S_{T_i}$, $P_{T_i-jj} = 1 - \sum_{k=0}^{n-1} p_{T_i-jk}, k \neq j$. G_{T_i} and S_{T_i} is the statistical average of the relevant data for this time period in consecutive days.

Set the transition probability sparse matrix at different time period in a day is P_i', $i \in (1,7)$, according to above, substituting P_{T_i-jk} and P_{T_i-jj} into P_i', the transition probability matrix of bike-sharing stations at T_i time period in a day is constructed .

Since P_i' is a random matrix, so 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). The rank-one updating method is first performed on the matrix Q, we get Q'. Based on Q', the steady-state probability vector and convergence time of the bike-sharing stations are 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 calculated separately.

The steady-state fleet size vector of bike-sharing stations at morning peak in a day is equal to the total size of the bicycles multiplied by the steady-state probability vector.

$$S_{T_1} = S * \pi \tag{9}$$

The steady-state fleet size vector for bike-sharing stations in the other time period in the same day is solved as follows

$$S_{T_i} = S_{T_{i-1}} P'_{i-1} = S_{T_{i-2}} P'_{i-2} P'_{i-1} = \cdots = S_{T_1} \left(\prod_{j=1}^{i-1} P'_j \right) \tag{10}$$

where S_{T_i} = the fleet size vector of bike-sharing stations at T_i time period in a day

P'_i = the transition probability matrix of bike-sharing stations at T_i time period in a day

7. CONCLUSION

In order to solve the problem of the fleet size and rebalancing of dockless bike-sharing stations, this paper improves the traditional Markov chain steady-state probability solving method, and proposes the “sparse matrix construction solution method”. The method constructs the transition probability sparse matrix combined with the practical application of dockless bike-sharing. Based on the analysis of reducible problems of the sparse matrix, the random prime matrix of transition probability is constructed to satisfy the requirements of steady-state fleet size vector calculation by rank-one updating method, and the convergence speed of the method is analysed. Thus, this paper further analyses the method of constructing the transition probability matrix in multiple time periods in one day, which makes the algorithm to solve the steady-state probability vector in different time periods in a day. Finally, the algorithm is verified with the practical and simulation data.

This algorithm takes into account the accuracy and speed of calculation, and the result of the steady-state fleet size of bike-sharing station can be used as a baseline reference data to dynamically optimize the fleet size of bike-sharing station operated by bike-sharing companies.

REFERENCES

Ching, W.K. and Ng, M.K., 2006. Markov chains. *Models, algorithms and applications*, pp. 1-46.

DeMaio, P., 2009. Bike-sharing: History, impacts, models of provision, and future. *Journal of public transportatio*n, 12(4), 3.

GuDong, 2017. From the ride data report by GuDong, we can see that small secrets of bike-sharing. http://www.ichuangye.cn/index.php/article-edit-2569-21.html Accessed 2 February 2018.

Hu, J., 2018. *User behaviour analysis and optimization in Bike-sharing systems*. Zhejiang University.

Jiao, G. Z., 2017. A brief analysis of the sustainable profit model of the Bike-Sharings. *Consume Guide*, 39(2), pp. 37-38

Liu, C. H., 2008. *Stichastic process (4th edition)*. Huazhong University of Science and Technology Press, pp. 65-72.

Langville, A.N. and Meyer, C.D., 2011. *Google's PageRank and beyond: The science of search engine rankings*. Princeton University Press, pp. 20-46.

Liu, C. J., 2012. Some Studies on the Calculation of Markov Transfer Matrix. *China external education*, (18), pp. 89-89.

MetroBike, 2018. bike-sharing. http:// bike-sharing.blogspot.com/ Accessed 2 April 2018.

Mobike, 2018a. RECYCLING, REFUNDING AND REDESIGNING: Unveiling Mobike's New Growth Initiatives. https://mobike.com/global/blog/post/growth-initiatives Accessed 7 July 2018.

Mobike, 2018b. Mobike About. https://mobike.com/global/about Accessed 2 April 2018.

Shao, D., and XUE, M. G., 2017. Bike-sharing and Sustainable Urban Development - China urban transport development BBS 2017 first symposium. *Urban Transport of China*, 15(03), pp. 1-6.

Stewart, W. J., 1994. *Introduction to the numerical solution of Markov chains*. Princeton University Press.

wikipedia, 2018. Dockless bikes. https://en.wikipedia.org/wiki/Bicycle-sharing_system#cite_note-21/ Accessed 9 May 2018.

Wang, G.R., 2017. The traffic value and development path of bike-sharing. *City*, 205(04), pp. 72-75.

Yang, Y. Q., 2017. Design and research on Crime Prevention Through Environmental about Bike-sharing. *Chinese Criminology Review*, 009 (2), pp. 67-74.