

## Research Article

# Assessing Accessibility of Dockless Sharing-Bike Networks by the Social Network Analysis Method

Pei Liu <sup>1</sup>, Junlan Chen <sup>1</sup>, Heyang Sun <sup>1</sup>, Xiucheng Guo <sup>1</sup>, Yan Wang <sup>2</sup>,  
and Zhenjun Zhu <sup>3</sup>

<sup>1</sup>School of Transportation, Southeast University, No. 2 Dongnandaxue Road, Nanjing 211189, China

<sup>2</sup>Chien-Shiung Wu College of SEU, Southeast University, No. 2 Dongnandaxue Road, Nanjing 211189, China

<sup>3</sup>College of Automobile and Traffic Engineering, Nanjing Forestry University, No. 159 Longpan Road, Nanjing 211189, China

Correspondence should be addressed to Xiucheng Guo; seuguo@163.com

Received 26 February 2021; Accepted 21 June 2021; Published 1 July 2021

Academic Editor: Luigi Dell'Olio

Copyright © 2021 Pei Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Dockless sharing bikes play an increasingly significant role in transit transfer, especially for the first/last mile. However, it is not always accessible for users to find sharing bicycles. The objective of this paper is to assess the accessibility of dockless sharing bikes from a network perspective, which would provide a decision-making basis not only for potential bike users but also for urban planners, policymakers, and bicycle suppliers to optimize sharing-bike systems. Considering bicycle travel characteristics, a hierarchical clustering algorithm was applied to construct the dockless sharing-bike network. The social network analysis (SNA) method was adopted to assess the accessibility of the bike network. Then, a spatial interaction model was chosen to conduct a correlation analysis to compare the accessibility obtained from the SNA approach. The case study of Shanghai indicates a strong connection between the accessibility and the SNA indicators with the correlation coefficient of 0.779, which demonstrates the feasibility of the proposed method. This paper contributes to a deep understanding of dockless sharing-bike network accessibility since the SNA approach considers both the interaction barriers and the network structure of a bicycle network. The developed methodology requires fewer data and is easy to operate. Thus, it can serve as a tool to facilitate the smart management of sharing bikes for improving a sustainable transportation system.

## 1. Introduction

Dockless bike-sharing has been launched in many counties worldwide since 2015. A still-growing list of cities that provide such service can be found at the bike-sharing world map. Dockless bike-sharing has sprung up in China since 2016, and the scale of shared bicycle users in China reached 235 million in 2018, indicating its extensive usage among urban transportation. The dockless sharing bikes are one kind of public transportation modes with unique strengths. However, there is still some room for improvements reflected in the dockless sharing-bike satisfaction survey. The biggest problem is that when passengers want to use sharing bikes to travel, it is difficult to find sharing bikes within the acceptable range. This indicates that it is extremely important for dockless sharing bikes to be accessible to the potential users, as accessibility is a

key indicator to evaluate the effective functioning of dockless sharing bikes. Therefore, to improve the satisfaction of dockless bike-sharing service, the accessibility of dockless sharing bikes should be assessed first to inform both suppliers and users.

This paper focuses on measuring the accessibility of dockless sharing bikes in a network perspective within relatively available data. Furthermore, we apply the social network analysis approach to the evaluation of the bicycle network, and a correlation analysis between accessibility and SNA indicators is conducted to verify the effectiveness and reliability of the proposed method. According to these research studies, the characteristics of accessibility can be summarized as follows: accessibility is a spatial-temporal concept, which represents the difficulty extent of communication between spatial entities to overcome the resistance of interactions. Accessibility includes three essential

elements: origins, destinations, and the transportation system. This research could help city planners, policymakers, and bicycle suppliers obtain information to identify areas that need to improve infrastructure investment and bike-sharing to create a friendly environment for nonmotorized travel patterns and improve sustainable transport systems. This is in line with the concept of green and sustainable development.

The rest of the paper is organized as follows: the next section is a literature review. Section 3 presents the methodology, including dockless sharing-bike network generation by a new clustering algorithm and accessibility assessment model construction based on the SNA approach. In Section 4, the data availability is explained and a correlation analysis is performed to explore the relationship between SNA indicators and accessibility. Finally, the paper finishes with some conclusions and proposals.

## 2. Literature Review

Bike-sharing allows people a healthy, enjoyable, and emission-free way to commute across small distances free from the worries of owning a bike [1], which is the primary motivation for sharing-bike usage [2, 3]. Particularly, dockless sharing bikes could be used in addition to public transit, thus providing an efficient option for first-/last-mile trips [4]. Integrating bike-sharing into urban transport systems is critical to mutually reinforcing sustainable transport networks [5]. In order to enable people to use shared bikes more, the improvement of accessibility is an important aspect. In this way, users can be notified of the best pick-up location in advance, and smart management as well as improvement of bike-sharing service can be realized by urban planners, policymakers, and suppliers [6, 7].

Defining and evaluating accessibility allows us to think about it systematically, as it relates to planning, transportation, equity, and other aspects [8]. Accessibility (or just access) for transportation planning is broadly referred to as the ease of reaching goods, services, activities, and destinations, which together are called opportunities [9]. Weibull defines accessibility as the freedom of individuals to participate in activities [10]. The notion of accessibility relies on two core concepts: location (transport infrastructures, population, and economic activity) and distance. Accessibility therefore takes into account a number of development patterns, including density and land use [11]. Murray and Wu believed that accessibility comprised access and geographic coverage in public transit planning. Access is the process of getting users from their origin to the transportation network, and geographic coverage is the transportation network across a spatial medium [12]. Zuo et al. defined bicycle-transit accessibility as the number of locations reachable within a time threshold [13].

In terms of quantitative evaluation of accessibility, numerous methods have been developed. Researchers have proposed several criteria to assess the usefulness and limitations of accessibility measures for different study purposes [14], such as reasonability, feasibility, interpretability, and usability [15]. The commonly used methods for accessibility measurement

are spatial barrier model, cumulative opportunity model, spatial interaction model, utility model, and space-time constraint model [16–19]. The spatial interaction model and utility model comprehensively consider land use and traffic demand, spatial barrier, and other factors, which can be widely used in the study of coordinated development of land use and traffic. The cumulative opportunity model is challenging to determine the time threshold, and its application scope is relatively limited, so it is mostly used in the project location, economic center node evaluation. The spatial barrier model and spatial-temporal constraint model cannot reflect other information such as land use and traffic demand, and the amount of data needed is large and difficult to obtain, so the application scope is limited.

Apart from these widespread models, some concepts from social network analysis (SNA) also have been introduced to quantify accessibility. SNA, also known as structural analysis, is a strategy for investigating social networks and structures using mathematical concepts abstracted from graph theory [20]. Since the first appearance in 1960, SNA has been applied extensively in many social science areas. In the field transportation, SNA techniques are first introduced to the business logistics and transportation community to study the dynamic flows of communication between members of a social network [21]. Gehrke et al. believe that network connectivity can be used as a measure of accessibility [22]. Graph theory, a useful tool of network analysis, is a branch of mathematics with broad and practical applications, which can also be used to model transportation systems [23]. Batty introduced graph theory to measure accessibility with respect to the physical transportation infrastructure [24]. Chen et al. selected several measures of the network topological structure to examine the accessibility of the metro network in Guangzhou [25]. More recently, Rubulotta et al. made comparisons between connectivity and traditional traffic planning measurement methods, and the results demonstrated the correlation between SNA centrality calculation and traffic accessibility and closeness [26]. El-Adaway et al. used SNA to analyze transportation networks, which proves its effectiveness as a supplementary tool for improving transportation planning [27]. Sarlas et al. proposed a new method by combining the indicator of betweenness in social networks with geographical accessibility to estimate real-life urban network accessibility [28]. Chen and Chang illustrated the relationship between connectivity and accessibility and proved that SNA is valid for connectivity assessment [29]. From all the research studies, we could find a new perspective, that is, the SNA approach, to assess transportation networks and detect the interconnectivity in transport networks. Conventional assessment methods evaluate individuals and attributes in an isolated manner, whereas SNA methods evaluate individuals and their relationships in the overall network, thus extracting the network's holistic patterns [30].

Generally, most previous research studies have been conducted on accessibility, although there still exist some research gaps. First, there are no best methods for accessibility assessment. Among various accessibility assessment methods, challenges are noted arising primarily from problems such as data acquisition, data quality, determination of model parameters, and network attributes

ignorance. Different from traditional research, the SNA method emphasizes on structural features of networks rather than on length and path within networks, which provides us a new overview in accessibility assessment. What's more, previous studies paid more attention to motorized modes, which could not apply to bicycles directly. Dockless sharing bicycles are frequently chosen as an addition to public transportation for the first/last mile. Thus, the trips are mostly short distance. Consequently, the network of dockless sharing bikes is connected in a local rather than a global manner compared with motorized networks. As stated by Ting Zuo, "the access to transit by bicycling is greatly determined by bike network connectivity" [31].

To fill these knowledge gaps, we use the SNA method to emphasize the structural characteristics of the network and consider more local connections in the study of accessibility. This article defines the network accessibility of dockless sharing bicycles as the difficulty of using sharing bikes to flow between origins and destinations in the network.

### 3. Methods

A network comprises nodes (vertices) and links (edges). The dockless sharing bikes have no fixed locations, and they could be parked anywhere, resulting in the difficulty in node determination. To abstract the dockless sharing-bicycle network, a novel clustering algorithm based on riding distance was proposed to generate nodes in a network, and the links were defined according to both the riding distance and the bike volume between two nodes. Based on the network, we constructed a model to assess accessibility stemmed from the SNA approach. To verify the proposed approach, we first calculated the accessibility with the commonly used method (spatial interaction model, Section 3.1) and compared the results of the two methods to check the efficiency of the new approach.

**3.1. Spatial Interaction Model.** In order to verify the effectiveness of the SNA method, this paper introduced the spatial interaction model as a control. In this model, accessibility was defined as the potential for interaction between two points, indicating the impedance between nodes and affected by the scale of traffic activities on the network. The calculation formula was based on the potential model proposed by Hansen [19]:

$$\text{Acc}_i = \sum_j \frac{D_j}{d_{ij}^\alpha}, \quad (1)$$

where  $\text{Acc}_i$  represents the accessibility of node  $i$ ,  $d_{ij}$  represents the impedance between two nodes in the network, and  $D_j$  is the number of opportunities provided by node  $j$  within a specific impedance acceptable range. In this article, impedance was measured by the nonlinear distance between network nodes, where  $D_j$  is the number of bicycles provided (output) by node  $j$  (in other words, the number of bicycles that start from node  $j$ ),  $d_{ij}$  represents the impedance between two nodes in the network, and  $\alpha$  is a parameter reflecting the degree of impedance influence, which could be selected according to the empirical value.

### 3.2. Accessibility Model Based on Social Network Analysis

**3.2.1. Node Generation.** The registration points of dockless sharing bikes present significant discontinuity and agglomeration in space. A hierarchical clustering method based on spatial distance was applied in this paper for two main reasons: the original dataset is too large to analyze as every bike could be a point on the map, and sharing bikes usually accumulate around POIs. Therefore, it will be convenient and reasonable to aggregate close locations in the same class. When it comes to the clustering criteria, the spatial distance was selected considering characteristics of dockless sharing bikes. Bike users generally choose dockless sharing bicycles to travel within one kilometer (seen in Figure 1), and they are usually not willing to walk long distances to find a dockless sharing bike. Distance is the primary factor that users might consider in bike usage. In this way, the spatial distance was selected as the standard for the clustering of origins and destinations, and the criteria scope should be constrained within the range of passengers' acceptable walking distances.

The iterative procedure of the distance-based clustering algorithm is presented in Figure 2, which combines the location of public transportation stations. Different from other clustering methods, there is no need to assign an exact number of final clusters in this paper. Circles with the same color denote bike origins and destinations that belong to the same group. Figure 2 represents the traversal order of these points. In the first iteration, each point did not belong to a cluster since no clusters had been defined. In every iteration, the algorithm searched for a point and determined if it belonged to an existing group based on the distance threshold. When the distance between this point and any center of existing clusters was less than the distance threshold, it would be put in this cluster set, and the center point coordinates would be updated. Otherwise, the algorithm created a new group whose center was the same as this point. After iterations, each point would be defined in a cluster that consisted of some spatially related points.

**3.2.2. Network Construction.** The travel of dockless sharing bikes presents evident uneven and geographically aggregative distribution in spatial as bicycles will gather and flow between different zones spontaneously. To be simplified, we assumed the clustering centers to be network nodes. After counting bike flow between clustering centers, reasonable flow impedance function could be set up to calibrate links in the network. In this paper, a bicycle network based on flow relationships and distances between different traffic zones was established:

$$G = (V, E), \quad (2)$$

where  $V$  represents the center nodes of dockless sharing-bike traffic zone, and  $E$  represents the link of nonmotorized traffic network, that is, the connection between nodes, which reflects the connection between traffic zones.

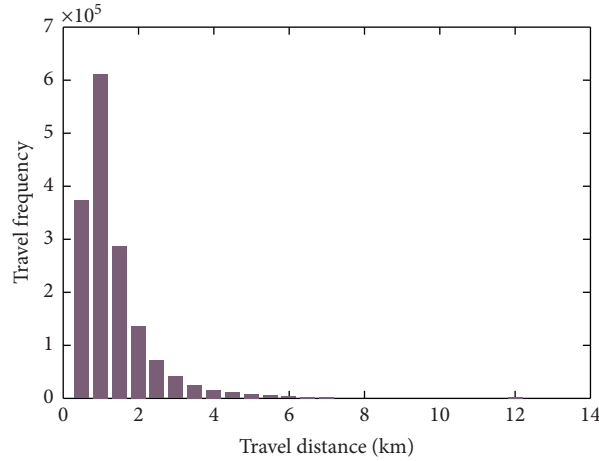


FIGURE 1: Distribution of dockless sharing bikes: travel distance distribution.

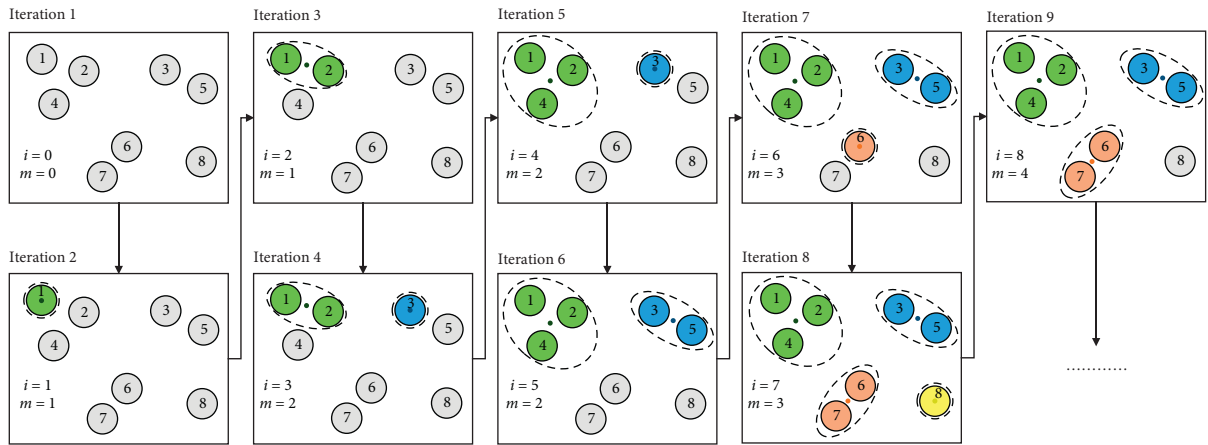


FIGURE 2: Iterative procedure of the distance-based clustering algorithm.

Since the traffic flow between two zones is usually asymmetric, in other words, there is commonly a significant flow direction between the two zones. Thus,  $G$  is actually a directed weighted graph. There are several unique features of a dockless sharing-bike network. First of all, a network is considered to be connected globally. In contrast, the dockless sharing-bike network is locally connected as it consists of multiple connected subgraphs responding to short-distance travel. Thus, the nodes far away from others are not connected. Moreover, the weight of links in this network is calibrated for accessibility assessment, which differs from that of conventional automobile networks. The higher the bike flow between two nodes is, the stronger the connection will be, while the longer the distance is, the weaker the link will be. And the weight will be infinitely great if the distance between two nodes exceeds the threshold of travel distance. According to the characteristics of a dockless sharing-bike network, the weight of links should be considered in two facets. In this paper, the weight of links was calculated as follows:

$$\text{weight} = \begin{cases} \frac{d}{q}, & d \leq d_0, \\ \infty, & d > d_0, \end{cases} \quad (3)$$

where  $d$  denotes the nonlinear distance of two nodes, which equals to Euclidean distance multiples nonlinear parameter (which is 1.3 according to the Code for Design of Urban Road Engineering (CJJ37-2012));  $d_0$  is the threshold of dockless sharing-bike travel distance (6 km); and  $q$  denotes the bicycle flow of a link.

**3.2.3. Model Index Selection.** There are several reasons for social network analysis to be applied to dockless sharing-bike accessibility assessment. First, substantial similarities could be observed between dockless sharing-bike networks and social networks from a structural perspective. The three components of a social network are actors, ties, and boundaries. Similarly, the three elements also exist in the network of dockless sharing

bikes in the form of nodes (cluster centers), links, and boundaries of research regions. Second, the purpose of social network analysis and accessibility analysis is consistent. The most prominent objective of the two studies is to discover the interdependence and connections between individuals (nodes) and display these connections in a network. Therefore, we assumed that it is reasonable to apply the approach of SNA in the accessibility analysis of a dockless sharing-bike network.

In social network analysis, different measures of centrality are commonly introduced to reflect the degree of connectivity of nodes in the network. Some scholars have already applied the SNA method in transportation network analysis, and “betweenness,” “connectivity,” and “centrality” are frequently selected as measurements. In this paper, the accessibility of a dockless sharing-bicycle network is defined as the difficulty of using sharing bikes to flow between origins and destinations in the network, which could be evaluated when the following questions are addressed:

- (i) To what extent is a node connected to other nodes?
- (ii) How far is the distance between the connected nodes?
- (iii) What is the role of the node in a network (a hub to connect other nodes or an endpoint)?

On this basis, we chose the following five indexes to reflect the characteristics of a sharing-bike network. The first two indexes deal with the first question, while the second and the last could be addressed by index 3, 4, and 5 independently.

(1) *Out-Degree Centrality*. The number of links out of the node in the directed graph, which represents the expansion force of the node, also known as integration.

$$D_i^{\text{out}} = \sum_{j \neq i \in E} \text{out}_{i,j}, \quad (4)$$

where  $j$  represents any node except the node  $i$ ,  $E$  represents the set of all network nodes, and  $\text{out}_{i,j}$  represents whether there is a reachable edge from  $i$  to  $j$ , when it is yes,  $\text{out}_{i,j} = 1$ , otherwise,  $\text{out}_{i,j} = 0$ .

(2) *In-Degree Centrality*. The number of links entering the node in the directed graph, showing the attraction of a node, or being expressed as radiality.

$$D_i^{\text{in}} = \sum_{j \neq i \in E} \text{in}_{i,j}, \quad (5)$$

when there is a reachable edge from  $j$  to  $i$ ,  $\text{in}_{i,j} = 1$ ; otherwise,  $\text{in}_{i,j} = 0$ .

(3) *In-Closeness Centrality*. The reciprocal of the sum of the shortest distances from a node to all other nodes, showing the difficulty of a node to other nodes.

$$C_i^{\text{in}} = \frac{1}{\sum_{j=1}^n d(i,j)}, \quad (6)$$

where  $d(i,j)$  represents the length of the shortest path from  $i$  to  $j$ .

(4) *Out-Closeness Centrality*. The reciprocal of the sum of the shortest distances from all other nodes to the node, showing how easy it is for other nodes to reach the node.

$$C_i^{\text{out}} = \frac{1}{\sum_{j=1}^n d(j,i)}, \quad (7)$$

where  $d(j,i)$  represents the length of the shortest path from  $j$  to  $i$ .

(5) *Betweenness Centrality*. The ratio of the shortest path passing through this node to all shortest paths, which reflects the degree of the close connection between this node and other nodes (Equation 8).

$$B_k = \sum_{(i,j) \in E} \frac{P_k(i,j)/P(i,j)}{(n-1)(n-2)/2}, \quad (8)$$

where  $P(i,j)$  represents the shortest paths between node  $i$  and  $j$ ,  $P_k(i,j)$  denotes the number of shortest paths between node  $i$  and  $j$ , including node  $k$ , and  $n$  is the number of nodes in the network. Visually, it can be found that the betweenness centrality reflects the importance of the node as an intermediate bridge, which indicates the degree of closeness.

The shortest path algorithm (Dijkstra algorithm) was involved in the index calculation. We determined a threshold of the shortest path solution, considering the characteristics of dockless sharing-bicycle travel. Then, the nonlinear distance between the origins and destinations was calculated based on the shortest path. When the result was larger than the threshold value, the path would be invalid. The threshold was set to be 6 km, as shown in Figure 1.

3.2.4. *Calculation of Accessibility*. The five indexes demonstrate the centrality of every node conclusively and exert different impacts on accessibility. In order to facilitate comparison, it is necessary to fuse the five indicators into one indicator as a measure of accessibility. We simplified the accessibility model that was quantitatively expressed as follows:

$$\text{Acc}_i = \sum_{j=1}^5 w_j C_{ij}, \quad (9)$$

where  $\text{Acc}_i$  denotes the accessibility of sharing-bike clustering center  $i$ ,  $w_j$  denotes the weight of the  $j$ th indicator, and  $C_{ij}$  denotes the  $j$ th indicator value of sharing-bike clustering center  $i$ .

The entropy method is usually employed to evaluate the influence of various indicators on comprehensive measurement. Information theory claims that the value of entropy is a measure of uncertainty. The entropy declines with the increasing amount of information, as well as the dispersion degree of an index. Therefore, the entropy method was selected in this paper to determine the weight of each indicator (coefficient  $w_j$ ). The calculation process consisted of three steps.

(1) *Normalization of Indicators*. It is necessary to standardize the original data as the units and dimensions of various indexes are of differences. Moreover, the order of magnitude of each index differs considerably. The maximum-minimum standardization method was used to convert the initial measurement values into nonunit variables (equation (10)).

$$E_i^j = \frac{X_i^j - \min X^j}{\max X^j - \min X^j}, \quad (10)$$

where  $E_i^j$  is the standardized value of the index  $j$  of research unit  $i$ ,  $X_i^j$  is the initial value of the index  $j$  of research unit  $i$ ,  $\min X^j$  is the minimum value of the index  $j$  of all research units, and  $\max X^j$  is the maximum value of the index  $j$  of all research units.

(2) *Index Entropy Calculation*

$$\delta_i^j = \frac{E_i^j}{\sum_{i=1}^n E_i^j}, \quad (11)$$

$$e_j = -\frac{1}{\ln n} \sum_{i=1}^n \delta_i^j \ln \delta_i^j,$$

where  $\delta_i^j$  is the proportion of the index  $j$  in the research unit  $i$ ,  $n$  denotes the number of research units, and  $e_j$  represents the entropy value of index  $j$ .

(3) *Weight Calculation of Indicators*

$$w_j = \frac{1 - e_j}{m - \sum_{j=1}^m e_j}, \quad (12)$$

where  $w_j$  is the weight value of indicator  $j$ , and  $m$  represents the number of indicators.

## 4. Case Study

4.1. *Data Availability*. Shanghai is an international metropolis where dockless sharing bikes gain broad popularity, and the network of sharing bikes is relatively stable. Therefore, we chose Shanghai as a case study. Data used in this paper consisted of three types: dockless sharing-bike dataset of Shanghai, Shanghai administrative division data, and points of interest (POI) data.

### 4.1.1. Data Sources

(1) *Dockless Sharing-Bike Dataset*. The dockless sharing-bike data used in this paper have approximately 20 million pieces of record in 14 days from August to September 2018 in Shanghai after decryption. The single-day data include bike ID, data time, lock status, and location (longitude and latitude) records of all dockless sharing bicycles in Shanghai.

Bike identification is used to identify bikes, and then the single-day running data record chain of a bicycle can be detected. Lock status is the state quantity of shared bicycles, which indicates whether the bike is in usage. Data generation times, as well as latitudes and longitudes, provide spatial and

temporal information of dockless sharing bicycles, which makes it possible to analyze the trajectory of dockless sharing bicycles quantitatively.

Due to the error of transmitting signal equipment of dockless sharing bikes, there may be time vacancies in the DATA\_TIME field, and these records need to be removed to ensure information integrity in subsequent processing. Furthermore, the latitude and longitude data of sharing bicycles may drift away from the actual data points because of the errors of the GPS. This could be solved by removing the records whose longitude and latitude are not located in the Shanghai administrative region.

(2) *Points of Interest (POIs) Data*. POI data are obtained from Baidu Map. There are approximately 550,000 POI records in Shanghai, which could be divided into 12 categories, containing catering, landscapes, public facilities, transportation facilities, education, finance districts, business residence, automobile services, sports, leisure, medical services, and accommodation. POIs could be used to verify the reliability of node generation.

(3) *Shanghai Administrative Division Data*. The data of Shanghai administrative divisions in Esri Shapefile format for 2018 include the boundary data of 16 municipal districts in Shanghai. This kind of data could be used in determining network boundaries.

4.1.2. *Characteristic of Dockless Sharing Bikes*. The characteristics and distribution of dockless sharing-bicycle trips are shown in Figure 3. Figure 3(a) indicates that the travel distance of sharing bicycles is mainly within 3 km (95.2%) and will not exceed 6 km. The highest frequency of travel distance ranks between 0.5 and 1.0 km, accounting for 38.3% of total trips. Figure 3(b) shows that 85.5% of the bike travel time is within 20 minutes, of which 35.3% is 5–10 minutes, corresponding to the travel distance of 0.5–1.0 km. It can be concluded that sharing bicycles provide a new choice for the first/last mile of urban travel, and most users choose sharing bikes for short-distance trips. In this case, sharing bikes are an essential part of the connection transportation method of subway and buses. The speed of sharing bicycles in Figure 3(c) approximately obeys a normal distribution with the mode of 7–8 km/h and relatively small discreteness. From Figure 3(d), a bimodal distribution of sharing bikes could be observed. The peak time of sharing-bike trips is at 9 a.m. and 7 p.m., which coincides with the peak time of commuting trips. It is worth mentioning that the duration of the morning peak and evening peak is different. The evening peak of sharing-bicycle usage is about three hours, while the morning peak is only one hour. This also reflects the diversity of get-off work hours, and the freedom of travel in the evening is relatively higher than morning. At the same time, the usage of sharing bicycles in the morning peak is significantly lower than that in the evening peak. This may be due to people's higher requirements for time reliability when traveling in the morning. There is a certain degree of randomness in finding sharing bicycles, and people tend to

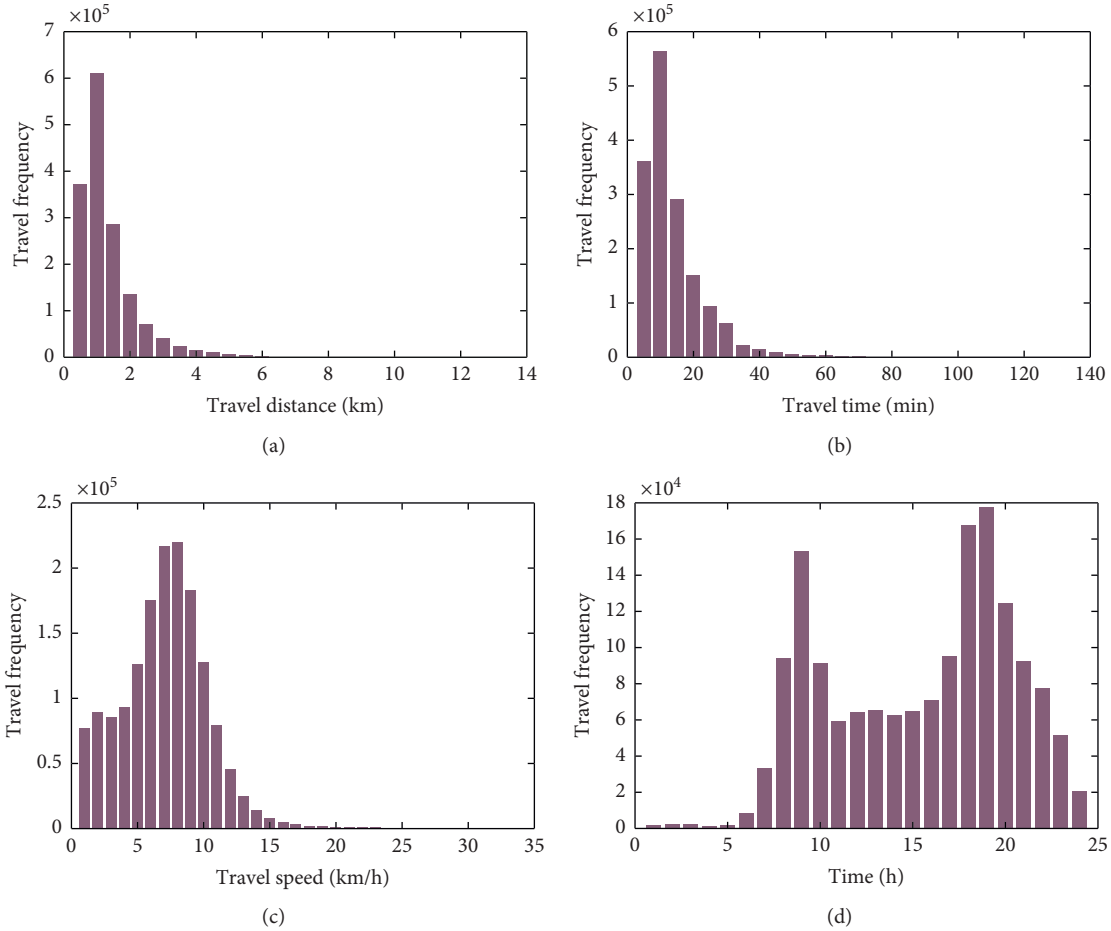


FIGURE 3: Distribution of dockless sharing bikes: (a) travel distance distribution, (b) travel time distribution, (c) travel speed distribution, and (d) 24-hour travel distribution.

choose more reliable travel modes to avoid being late. It can be inferred that most bicycle trips are commuting trips—connecting subways and buses in the first/last mile. These features of dockless sharing-bike travel can be applied to the calibration of model parameters.

**4.2. Results and Discussion.** The clustering algorithm was applied to cluster about 20 million origin and destination points of dockless sharing-bike travel records in Shanghai within two weeks, and approximately 8,000 sharing-bike traffic zones were obtained. The information of this network is displayed in Table 1. Each zone is a convex polygon with an area of about  $0.25 \text{ km}^2$  (Figure 4). Comparing these clustering centers with POIs, we could find that most of them are coincident with POIs, indicating that bikes spontaneously gather around POIs. The division of bike traffic zones not only lays a foundation for the follow-up research of this paper but also provides a new idea for dealing with the network distribution of nonmotorized traffic.

Based on the data of bike travel flow and travel district division obtained by clustering analysis, a sharing-bike network was established. Afterward, the SNA method was

applied to calculate the in-degree, out-degree, in-closeness centrality, out-closeness centrality, and betweenness centrality of each node before the entropy method was introduced to compute the weight of each index. On this basis, the accessibility of each sharing-bike traffic zone was obtained. The final model is in the format of equation (9) and could be expressed as equation (13). The in-closeness centrality and betweenness centrality indexes exerted a stronger influence on the accessibility, which demonstrated that the difficulty of reaching other nodes and closeness are the critical factors in accessibility assessment.

$$Acce_i = 0.160D_i^{\text{out}} + 0.172D_i^{\text{in}} + 0.208C_i^{\text{in}} + 0.200C_i^{\text{out}} + 0.260B_i. \quad (13)$$

The five indicators and accessibility values of each clustering zone based on the SNA approach are displayed in Figure 5. The five indicators show a decreasing trend from the city center to the periphery. In addition to the city center, there are other subcenters, which also have relatively high centrality and availability. Observing figures of the five indicators, we could discover that there was a big difference in the evaluation of the peripheral areas. The values of closeness centrality and betweenness centrality in the city periphery were almost close

TABLE 1: The information of the sharing-bike network.

Number of nodes	Number of edges	Mean closeness	Mean betweenness	Mean clustering coefficient
8061	106297	$3.205 \times 10^{-4}$	$7.179 \times 10^{-4}$	0.292

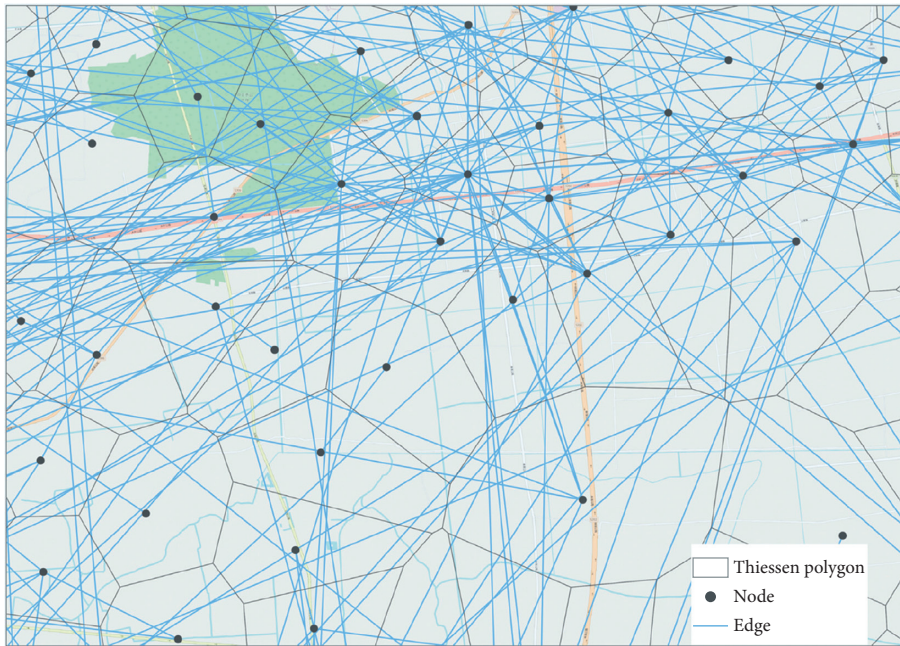


FIGURE 4: The sharing-bike network of Shanghai (part of the network).

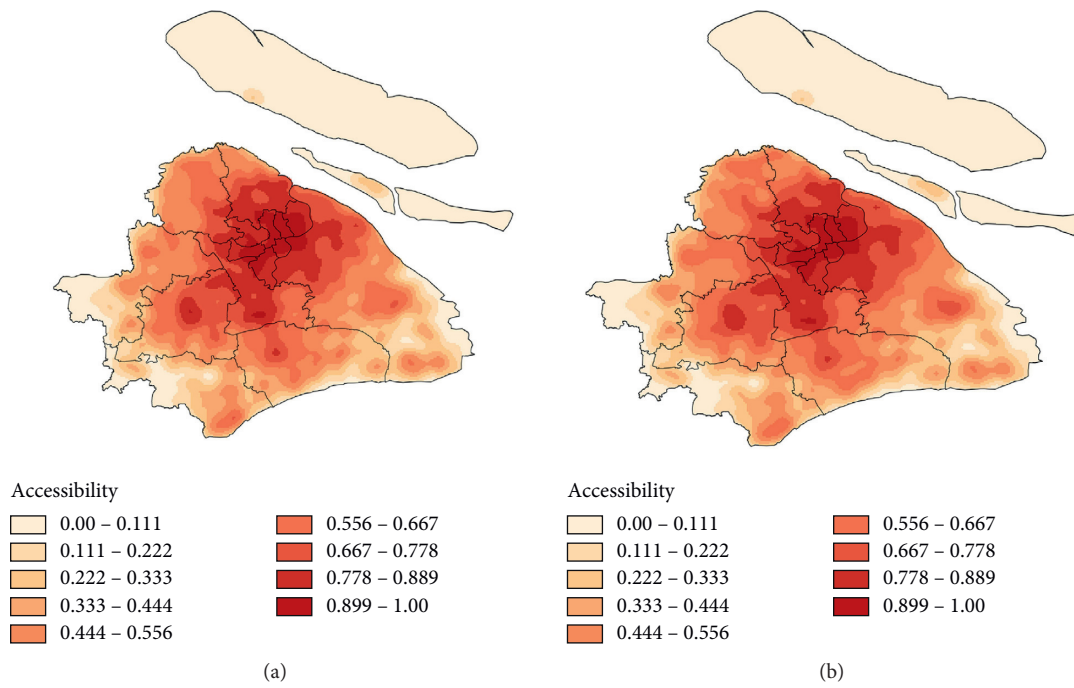


FIGURE 5: Continued.



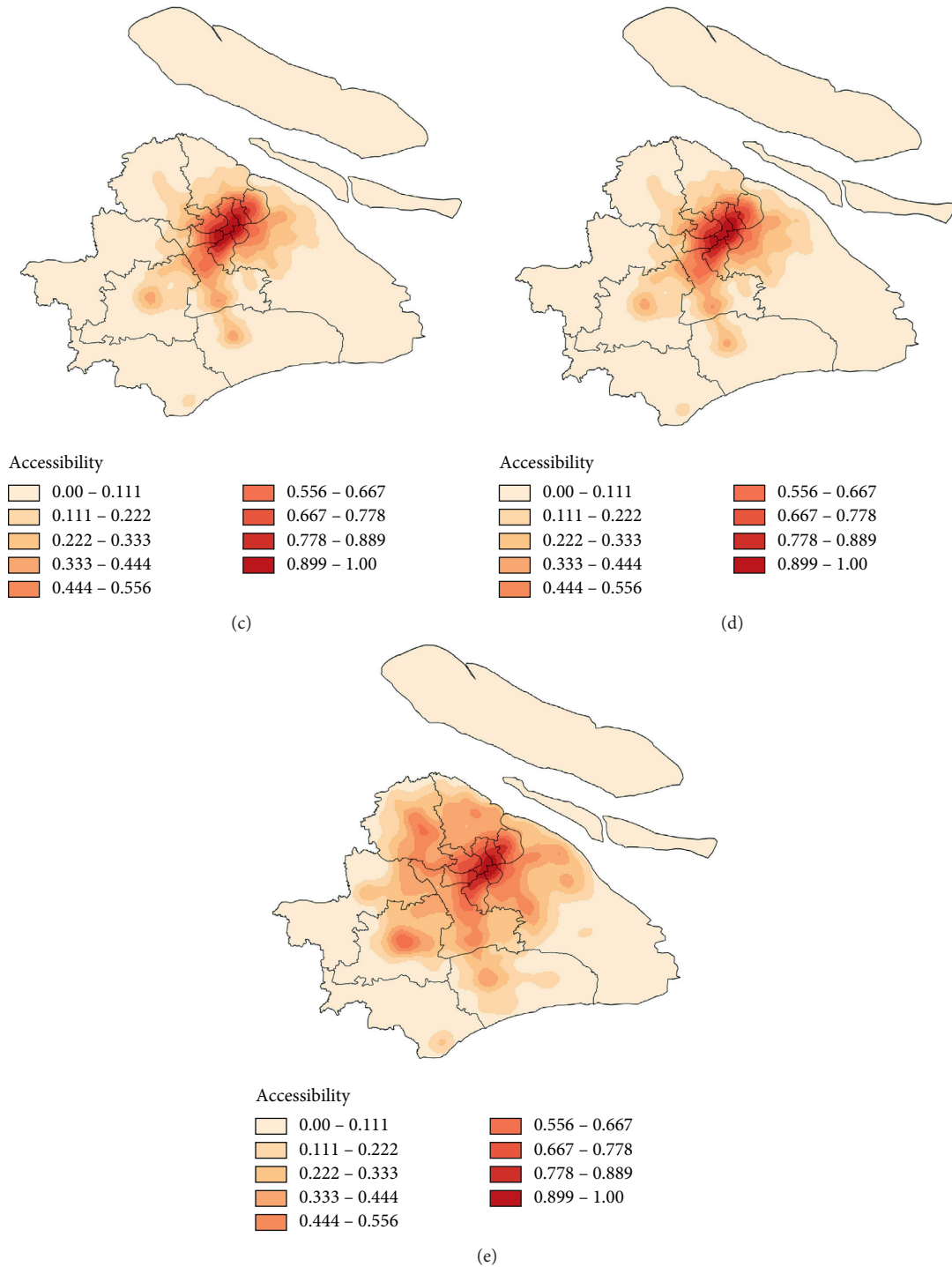


FIGURE 5: The distribution of indicators from SNA for the dockless sharing-bike network in Shanghai: (a) out-degree centrality distribution, (b) in-degree centrality distribution, (c) out-closeness centrality distribution, (d) in-closeness centrality distribution, and (e) betweenness centrality distribution.

to zero, which may result from the relatively small supply of bikes in the peripheral area and the longer distances between POIs. However, the connection function of nodes (the number of links of a node) in the network is ignored in those indexes. Therefore, the accessibility calculation model that integrates

various indicators can evaluate the accessibility of the entire research area from multiple angles.

To explore the relationship between different indicators, we conducted a Pearson correlation analysis (Table 2). As shown in the table, there is a relatively strong correlation

TABLE 2: Correlation matrix of accessibility measures.

	In-degree centrality	Out-degree centrality	In-closeness centrality	Out-closeness centrality	Betweenness centrality	Acc_SNA	Acc_SIM
In-degree centrality	1						
Out-degree centrality	0.642** (0.000)	1					
In-closeness centrality	0.302** (0.000)	0.328** (0.000)	1				
Out-closeness centrality	0.289** (0.000)	0.329** (0.000)	0.982** (0.000)	1			
Betweenness centrality	-0.023 (0.000)	-0.003 (0.000)	0.754** (0.000)	0.760** (0.000)	1		
Acc_SNA	0.544** (0.000)	0.571** (0.000)	0.929** (0.000)	0.928** (0.000)	0.721** (0.000)	1	
Acc_SIM	0.319** (0.000)	0.350** (0.000)	0.847** (0.000)	0.841** (0.000)	0.502** (0.000)	0.779** (0.000)	1

\*\*Correlation is significant at the 0.01 level (2-tailed). Acc\_SNA represents that the accessibility is obtained by the SNA approach, while Acc\_SIM represents that the accessibility is by the spatial interaction model.

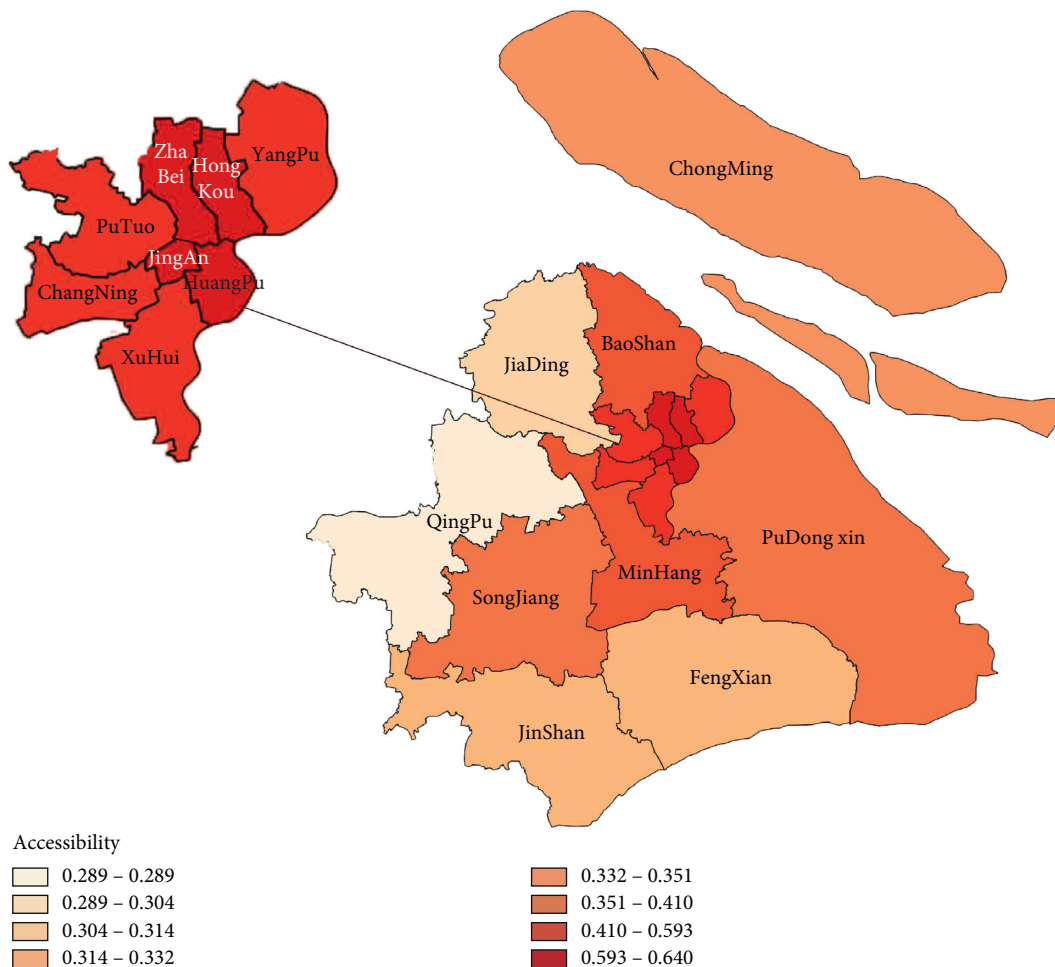


FIGURE 6: Accessibility map of dockless sharing bikes in each administrative region of Shanghai.

between accessibility (measured by the spatial interaction model) and the closeness centrality (correlation coefficient = 0.84). Accessibility (measured by spatial interaction model) is moderately correlated with betweenness centrality and poorly connected with node degrees. Since closeness

centrality is the reciprocal of the sum of the shortest distances between two nodes, the high correlation indicates that accessibility primarily deals with the travel barrier, in other words, distance plays a vital role in accessibility assessment. The entropy value of accessibility based on SNA shows a

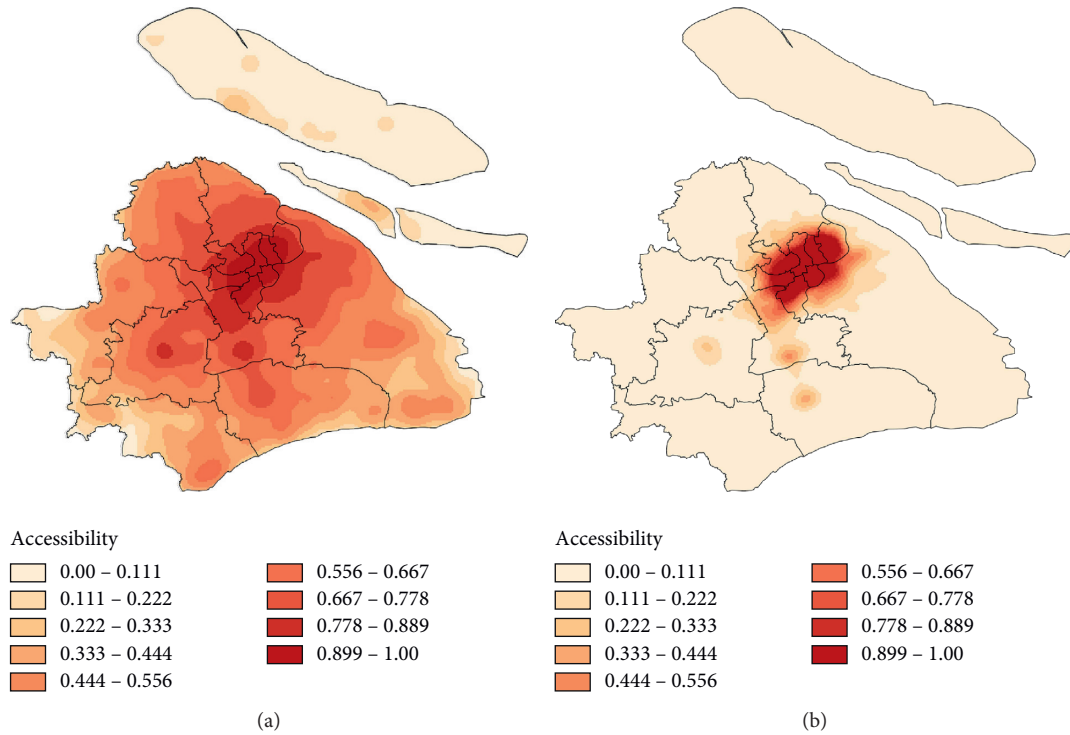


FIGURE 7: Accessibility map of the dockless sharing-bike network in Shanghai: (a) accessibility based on the SNA approach and (b) accessibility based on the spatial interaction model.

relatively high correlation with that from the spatial interaction model. In this way, we could believe that this proposed method could be complementary for accessibility assessment. These five indicators, as part of the accessibility assessment, provide a more comprehensive perspective for accessibility evaluation. Apart from the distance between origins and destinations, the new method also pays attention to the characteristics of the dockless sharing-bike network and considers the threshold of bicycle travel to determine whether the nodes are connected or not. Therefore, accessibility assessment based on the SNA approach could be more in line with reality.

To be more intuitive for decision-makers, we display the accessibility of an administrative region, which was defined in this paper as the summary of accessibility of each traffic zone divided by the total number of traffic zones in this administrative region. Then, the accessibility of each administrative region in Shanghai can be seen in Figure 6. From the figure, the accessibility of dockless sharing bikes in Jing'an District, Huangpu District, Hongkou District, and Minhang District ranked the highest, which meant that it was more convenient for passengers to use dockless sharing bicycles in these areas, that is, these areas were more friendly to nonmotorized traffic. However, peripheral administrative regions such as Jinshan District, Qingpu District, Jiading District, and Chongming District were less accessible, indicating that nonmotorized transport in these zones was underdeveloped. The results illustrated the unbalanced development of nonmotorized transportation, which was primarily because of differences in infrastructure supplies, bike numbers, and geographical locations. More research

could be conducted to find the principal causes based on accessibility results. Urban planners, policymakers, and bicycle suppliers could obtain information for optimization decisions.

**4.3. Model Evaluation.** The five indicators and accessibility values of each clustering zone based on the SNA approach are displayed in Figure 7(a). The accessibility measured by the spatial interaction model is displayed in Figure 7(b), which shows a similar distribution with closeness centrality. Its ability to evaluate the accessibility of the peripheral areas is weak since it cannot reflect the differences of the periphery. More abundant data may be required to support the study of the entire region.

## 5. Conclusions

This paper investigated the problem of dockless sharing-bike network accessibility from a network perspective. The SNA was applied to estimate the accessibility, considering the interaction barriers and the network structure of a bicycle network. The developed methodology requires fewer data and is easy to operate.

A dockless sharing-bike traffic network was founded on the basis of a hierarchical clustering algorithm according to the geographical correlation of origins and destinations. The algorithm was based on the characteristic of bike users, who would only be willing to find a sharing bike within his/her acceptable walking distance. Different from the  $k$ -means clustering algorithm, there is no need to assign a  $k$ -value in

advance, which otherwise would result in excessive clustering radius and inconsistency with user psychology. As a result, this algorithm clustered origins and destinations of bike travel from 20 million to approximately 8,000 sharing-bike traffic zones, which were regarded as nodes of the bike network. Compared with POI data, these zone centers almost overlapped with transit stops, schools, office buildings, shopping malls, etc., which indicated the reasonability of the algorithm.

An accessibility assessment model of sharing-bike network was constructed based on the SNA approach. Five indexes of social network analysis were introduced to evaluate accessibility, and the weight of each index was calculated according to their impacts on accessibility. This method considered not only the spatial barrier between origins and destinations but also focused on the node connectivity from a network perspective, thus providing a relatively comprehensive perception of sharing-bike network accessibility. The correlation analysis with the commonly used model (spatial interaction model) revealed that there exist some correlations between SNA indicators and accessibility. Therefore, it is reasonable to apply the SNA approach to assess the accessibility of sharing-bike network. The results suggested that the central districts like Jing'an, Huangpu, Hongkou, and Minhang District in Shanghai have higher sharing-bike accessibility than those peripheral administrative regions. Unlike other studies, the data required in this paper are easy to obtain as we could retrieve running data from bike-sharing companies to do real-time computing and present these results to both users and suppliers.

This paper assessed and mapped the accessibility of the dockless sharing-bike network in Shanghai, and the main work could be divided into two parts. A novel and promising perspective from SNA is proposed in accessibility assessment by analyzing the relationship between accessibility and different centralities. The objective of this paper is to provide information for the optimization of sharing-bike network, so accessibility was selected as an evaluation of a network current status. The results of accessibility could provide information for both users and decision-makers. The real-time assessment could be completed to display immediate accessibility for users to judge which place nearby is of higher accessibility. For urban planners, policymakers and bicycle suppliers could obtain information to identify areas where improving infrastructure investment and sharing bikes are of necessity, thus creating a user-friendly environment for nonmotorized travel mode and improving a sustainable transportation system. As the primary travel mode of nonmotorized traffic, sharing bikes are in line with the concept of green travel and sustainable development.

One of the limitations of this paper is that the data were collected in two weeks, which is not typical enough, and the estimation model cannot be extended to other seasons. In the future, it is essential to admit that one limitation of our approach is that the two-week data used in our research were not typical enough. The characteristics of cyclists may vary a little according to seasons, weathers, and holidays. Further investigations could be carried out to explore the thresholds used in different parts. Additionally, the proposed accessibility assessment model is mainly based on the

characteristics of network topology, and more concentration could be laid on the combination of different methods to understand and interpret the accessibility of a nonmotorized network thoroughly. Due to the limited data and time, the research does not focus on the other modes and types of networks to verify the extrapolation of the accessibility indicator weights. Based on collecting bike-sharing operation data in multiple cities, the future study can focus on extrapolating of the results.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

The authors acknowledge the data provided by Shanghai Big Data Joint Innovation Laboratory. Additional thanks go to their supervisor Prof. Xiucheng Guo and Meina Zheng for insightful comments and suggestions on the interpretation of the whole paper. This work was supported by the Scientific Research Foundation for Advanced Talents of Nanjing Forestry University (163106041), the General Project of Philosophy and Social Science Foundation of the Jiangsu Higher Education Institutions of China (2020SJA0125), and the General Program of the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (20KJB580013).

## References

- [1] A. Pal and Y. Zhang, "Free-floating bike sharing: solving real-life large-scale static rebalancing problems," *Transportation Research Part C*, vol. 80, 2017.
- [2] E. Fishman, "Bikeshare: a review of recent literature," *Transport Reviews*, vol. 36, no. 1, 2016.
- [3] Y. Ai, Z. Li, and M. Gan, "A solution to measure traveler's transfer tolerance for walking mode and dockless bike-sharing mode," *The Journal of Supercomputing*, vol. 75, no. 6, pp. 3140–3157, 2019.
- [4] Z. Chen, D. v. Lierop, and D. Ettema, "Dockless bike-sharing systems: what are the implications?" *Transport Reviews*, vol. 40, no. 3, 2020.
- [5] S. Katja and S. Joachim, "Spatial and temporal analysis of bike-sharing use in cologne taking into account a public transit disruption," *Journal of Transport Geography*, vol. 92, 2021.
- [6] A. Kaltenbrunner, R. Meza, J. Grivolla et al., "Urban cycles and mobility patterns: exploring and predicting trends in a bicycle-based public transport system," *Pervasive and Mobile Computing*, vol. 6, no. 4, 2010.
- [7] X. Zhou, "Understanding spatiotemporal patterns of biking behavior by analyzing massive bike sharing data in chicago," *PLoS One*, vol. 10, no. 10, 2015.

- [8] W. Manuel, "Taking one step further-advancing the measurement of green and blue area accessibility using spatial network analysis," *Ecological Indicators*, vol. 126, 2021.
- [9] T. Litman, "Evaluating accessibility for transportation planning," *Victoria Transport Policy Institute*, vol. 49, 2016.
- [10] J. W. Weibull, "An axiomatic approach to the measurement of accessibility," *Regional Science and Urban Economics*, vol. 6, no. 4, pp. 357–379, 1976.
- [11] J. Amila, P. K. S. Mahanama, and A. Chethika, "A modeling framework: to analyze the relationship between accessibility, land use and densities in urban areas," *Sustainability*, vol. 13, no. 2, 2021.
- [12] A. T. Murray and X. Wu, "Accessibility tradeoffs in public transit planning," *Journal of Geographical Systems*, vol. 5, no. 1, 2003.
- [13] T. Zuo, H. Wei, N. Chen et al., "First-and-last mile solution via bicycling to improving transit accessibility and advancing transportation equity," *Cities*, vol. 99, 2020.
- [14] M. Iacono, K. J. Krizek, and A. El-Geneidy, "Measuring non-motorized accessibility: issues, alternatives, and execution," *Journal of Transport Geography*, vol. 18, no. 1, 2009.
- [15] S. Zhang and Y. Zhang, *Analysis of Network Accessibility*, Springer, Cham, Switzerland, 2015.
- [16] K. T. Geurs and B. v. Wee, "Accessibility evaluation of land-use and transport strategies: review and research directions," *Journal of Transport Geography*, vol. 12, no. 2, 2003.
- [17] W. Martin and K. T. Gordon, "Physical accessibility as a social indicator," *Pergamon*, vol. 7, no. 5, 1973.
- [18] J. A. Pooler, "The use of spatial separation in the measurement of transportation accessibility," *Transportation Research Part A*, vol. 29, no. 6, 1995.
- [19] W. G. Hansen, "How accessibility shapes land use," *Journal of the American Planning Association*, vol. 25, no. 2, 1959.
- [20] E. Otte, "Social network analysis: a powerful strategy, also for the information sciences," *Journal of Information Science*, vol. 28, no. 6, 2002.
- [21] D. M. Phillips and J. K. Phillips, "A social network analysis of business logistics and transportation," *International Journal of Physical Distribution & Logistics Management*, vol. 28, no. 5, 1998.
- [22] S. R. Gehrke, A. Akhavan, P. G. Furth et al., "A cycling-focused accessibility tool to support regional bike network connectivity," *Transportation Research Part D*, vol. 85, 2020.
- [23] Y. Zhiyan, L. X. Cathy, M. Nikola et al., "Inferencing hourly traffic volume using data-driven machine learning and graph theory," *Computers, Environment and Urban Systems*, vol. 85, 2021.
- [24] M. Batty, "Accessibility: in search of a unified theory," *Environment and Planning B: Planning and Design*, vol. 36, no. 2, 2009.
- [25] S. Chen, C. Claramunt, and C. Ray, "A spatio-temporal modelling approach for the study of the connectivity and accessibility of the Guangzhou metropolitan network," *Journal of Transport Geography*, vol. 36, 2014.
- [26] E. Rubulotta, M. Ignaccolo, G. Inturri et al., "Accessibility and centrality for sustainable mobility: regional planning case study," *Journal of Urban Planning and Development*, vol. 139, no. 2, 2013.
- [27] I. H. El-Adaway, I. S. Abotaleb, and E. Vechan, "Social network analysis approach for improved transportation planning," *Journal of Infrastructure Systems*, vol. 23, 2016.
- [28] G. Sarlas, A. Páez, and K. W. Axhausen, "Betweenness-accessibility: estimating impacts of accessibility on networks," *Journal of Transport Geography*, vol. 84, 2020.
- [29] J. Chen and Z. Chang, "Rethinking urban green space accessibility: evaluating and optimizing public transportation system through social network analysis in megacities," *Landscape and Urban Planning*, vol. 143, 2015.
- [30] R. C. Hulst, "Introduction to social network analysis (SNA) as an investigative tool," *Trends in Organized Crime*, vol. 12, no. 2, 2009.
- [31] T. Zuo, H. Wei, and A. Rohne, "Determining transit service coverage by non-motorized accessibility to transit: case study of applying GPS data in Cincinnati metropolitan area," *Journal of Transport Geography*, vol. 67, 2018.