IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received September 13, 2018, accepted September 25, 2018, date of publication September 28, 2018, date of current version October 19, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2872451

Service Sites Selection for Shared Bicycles Based on the Location Data of Mobikes

BIAO JIN[1,2], MEI-HUANG HONG[3], XUE LIN[1], AND JIN-MING SHA[1]

[1]College of Geographical Sciences, Fujian Normal University, Fuzhou 350007, China
[2]College of Mathematics and Informatics, Fujian Normal University, Fuzhou 350108, China
[3]Fujian Newland Software Engineering Co., Ltd., Fuzhou 350015, China

Corresponding authors: Biao Jin (jinbiao@fjnu.edu.cn) and Jin-Ming Sha (jmsh64@sina.com)

This work was supported in part by the EU Eramus+ Project under Grant 586037-EPP-1-2017-1-HU-EPPKA2-CBHE-JP, in part by the project of the Fujian Provincial Department of Science and Technology under Grant 2018I0005, and in part by the project of the Fujian Provincial Department of Education under Grant JAT170149.

ABSTRACT Mobike is a leader in the bicycle-sharing industry, which commits to solving short-distance travel problems through the Internet. The rapid development of shared bicycles has been bringing very convenient for our travel. However, the problems of parking, high damage rates, and difficulty in reclaim have made many urban public space resources be taken up by these bicycles. In order to help the company maintain and manage its bicycles and service sites, the main contents of this paper focus on selecting the optimal service sites for Mobikes and planning the shortest circuit planning between service sites in each region. Our experimental data are the parking locations of Mobikes in Fuzhou city. The density-based clustering algorithm and an improved ant colony algorithm are adopted in this paper: 1) to divide a large data set into N small uncorrelate data sets, the density-based clustering algorithm reveals the high-density region by finding the low-density region of the dataset, and this type of algorithm is used for selecting the service sites and 2) the problem that starting from the warehouse center, then traversing all the service sites in the region, and finally returning to the warehouse center can be classified as an NP-hard problem. In this paper, an improved ant colony algorithm is used to obtain the optimal solution. The main purpose of finding the shortest circuit traversing all service sites in the region is to provide better support and maintenance for each service site.

INDEX TERMS Service sites, mobike, Fuzhou city, DBSCAN, ant colony algorithm.

I. INTRODUCTION

Our lives are changing in terms of sharing applications. In 2014, OFO and Mobike put forward the idea of using shared bicycles to solve the "last mile," thus opening up a new era of dockless bicycle sharing. Dockless bicycle sharing allow us to have access to an unprecedented means of transport. Since the bicycle-sharing market is characterized by a low entry threshold and easily replicated business mode, nearly 20 companies, including OFO, Mobike and U-bicycle, have entered the market in a short period of time (Figure 1).

These companies compete in the streets and alleys of various cities in mainland China. According to the 40th and 41st 'statistical reports on the internet development in China' issued by CNNIC (China Internet Network Information Center), from June 2017 to December 2017, the scale of bicycle-sharing users has increased from 106 million to 221 million. The rate of growth is up to 108.1%. At the same time, the utilization rate of shared bicycles has increased from

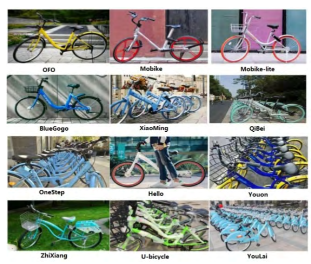


FIGURE 1. Main brands of sharing bicycles.

14.1% to 28.6% [1]. By the end of 2017, the business of bicycle-sharing had covered 21 overseas countries. The emergence of bicycle-sharing has changed many people's travel habits. They actively respond to green travel, making a great contribution to reducing both traffic congestion and carbon emissions. By now, the total number of cycling kilometers of bicycle-sharing users has exceeded 29.947 billion kilometers, thus reducing the carbon emissions by over 6.99 million tons. In regard to stimulating employment, the bicycle-sharing industry has created more than 30,000 off-line operation and maintenance posts [1]. Overall, the emergence of bicycle-sharing has greatly changed people's travel habits, reduced the congestion of urban traffic, and played an important role in promoting energy conservation, emission reduction, and green and sustainable urban development. However, what cannot be ignored is that they also causes a series of problems because of some subjective and objective factors.

At present, competition in bicycle-sharing industry is fierce. In order to occupy the market quickly, all companies launch their shared bicycles on a large scale. These bicycles have been taking up and encroaching upon limited city space. According to [2], in the past two years, the number of bicycle-sharing companies has reached 77, and a total of 23 million shared bicycles have been put into the market. Putting bicycle-sharing into the market without considering the actual demand has resulted in waste of resources. In some regions, the number of bicycles has far exceeded the demand. Many shared bicycles are idle at their parking points, leading to road traffic jams (Figure 2).

FIGURE 2. Shared bicycles occupying city roads.

Another bad consequence of blind delivery is that the supply of bicycles exceeds demand in some cities or regions, while others are in short supply. According to the relevant departments in Shanghai, as of the end of September 2017, there are now 1.15 million shared bicycles in the city [3]. The total number of shared bicycles in Guangzhou city of Guangdong Province is 8.65 million, while the delivery number in Dongguan city is only about 775,000 [4]. The proportion is seriously unbalanced. The number of shared bicycles in Wuhan city of Hubei Province is also up to 700,000. Additionally, there are no fixed parking points for dockless bicycles, which makes operators unable to monitor them in real time, and thus leaving them without any guaranteed maintenance or inspection. Although the government has issued guidelines and information for standardizing the operation of bicycle-sharing, there have been a wide variety of uncivilized uses of shared bicycles, such as adding private locks, violently dismantling locks, privatizing bicycles or even malicious discarding or incinerating bicycles (Figure 3).

Parking the shared bicycles in a disorderly manner has affected the appearance of the city. In view of these problems, our work mainly focus on the following two aspects: 1) Selecting service sites for the shared bicycles based on the distribution data of Mobikes in the five districts of Fuzhou city. The purpose of this is to prolong Mobikes' service life and reduce the operating cost of bicycle-sharing companies by providing quicker and more direct management and maintenance for Mobikes. 2) Planning the shortest circuit that can travel through all the service sites in each region respectively based on data about the selected service sites. The purpose of this is to provide better maintain and manage for the selected service sites themselves.

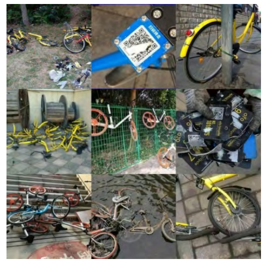


FIGURE 3. The broken sharing bicycles.

II. RELATED WORK

In order to accomplish our first goal, we need to conduct clustering analysis on the experimental data. According to the development of clustering algorithm, we can divide it into two categories: traditional clustering algorithms (including the Hierarchical methods, the Partition-based methods, the Density-based methods and the Grid-based methods) and modern clustering algorithms such as the fuzzy clustering. The comparison of these kinds of clustering algorithms is shown in the table below (Table 1).

TABLE 1. Comparison of clustering algorithms.

Algorithm type	Representative algorithms	Advantages	Disadvantages
Hierarchical methods	BIRCH Algorithm [5]	Can produce high quality clustering and solve the non spherical cluster problem	It has a high time complexity; It is difficult to deal with different sized clusters and convex shapes.
Partition-based methods	K-MEANS Algorithm [6]	It has a low time complexity and space complexity.	When the data set is very large, the result is easy to appear local optimum; It is very sensitive to noise; Can only be used for numerical type data.
Density-based methods	DBSCAN Algorithm [7]	It is not sensitive to noise; It can detect clusters of arbitrary shapes.	The result of clustering is strongly related to the value of parameters.
Grid-based methods	STING Algorithm [8]	The speed of clustering is fast because it is independent of the number of data objects.	It is sensitive to parameters and can not deal with irregular data. The accuracy of clustering results is low.
Fuzzy clustering method	Fuzzy C-means Clustering Algorithm [9]	For the data satisfying normal distribution, it will get very good clustering results. It is sensitive to isolated points.	The performance of the algorithm depends on the initial clustering center.

In view of the comparison of advantages and disadvantages between the above algorithms, and considering the characteristics of our experimental data (the bicycles are usually clustered in specific areas and times), we finally decided to adopt the Density-based methods for service sites selection.

DBSCAN (abbreviation of '*Density-Based Spatial Clustering of Applications with Noise*') algorithm was proposed in 1996 [7] and has become the representative of Density-based methods. Since it was proposed, there are more and more researchers have been concentrate their efforts on improving its performance or adopting it to solve various practical problems directly. Reference [10] concerned the parallelization of DBSCAN to improve its performance on high-dimensional data. To attack the slow speed of the AntClass algorithm (a new algorithm applying ant colony clustering algorithm to cluster analysis), [11] combined AntClass with DBSCAN and then proposed a new algorithm named 'DBAntCluster'. Reference [12] surveyed some important techniques in which original DBSCAN is modified or enhanced with improvement in complexity or result improvement on varied densities. In order to process the natural data which is often vague well, [13] presented an efficient clustering technique, named 'Soft DBSCAN' by combining DBSCAN with fuzzy set theory. Researchers of [14] survey over different variations of DBSCAN algorithms that were proposed as far as they know. They critically evaluated those variations and listed their limitations. Reference [15] modified the Imperialist Competitive Algorithm (ICA) with a density-based algorithm and fuzzy logic for optimum clustering in wireless sensor networks. Reference [16] presented a new algorithm based on DBSCAN named 'RDD-DBSCAN' to overcome the scalability limitations of the traditional DBSCAN algorithm by operating in a fully distributed fashion.

TABLE 2. Comparison of commonly used algorithms for solving TSP.

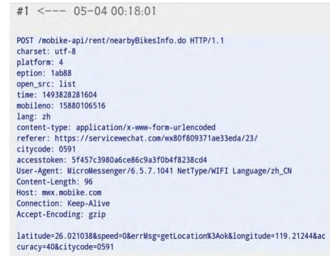
Algorithm type	Advantages	Disadvantages
Dynamic programming	It can find the optimum solution.	The computational complexity of this algorithm is high. It's not very suitable for solving TSP problem.
Branch-bound method	The optimal solution can be obtained and the average speed is fast.	It requires a large amount of memory spaces for running. Whether the algorithm is successful or not depends critically on the preset boundary value.
Genetic algorithm	It implies parallelism and global search. The optimal solution can be obtained with high running speed.	Its local search capability is bad.
Ant colony algorithm	The convergence rate of the optimal solution is fast and the result is stable.	The selection of pheromones and the lack of information hormones in its initial stage make the solution speed slow.

Essentially, the second research goal of this paper is a typical traveling salesman problem (TSP). As far as we know, the algorithms for solving TSP can be classified into two kinds. One is the traditional deterministic algorithm, such as dynamic programming [17] and branch-bound method [18]. The other is the modern intelligent algorithm, such as genetic algorithm [19] and ant colony algorithm [20]. The advantages and disadvantages of these commonly used algorithms are shown in the following table (Table 2). Considering the solution speed and stability of the optimal solution, this paper decides to adopt ant colony algorithm.

Many researchers have adopted ant colony algorithm to solve TSP problem [21]–[26]. By comparison, we can find that some researchers directly use the original ant colony algorithm [21] while others use an improved ant colony algorithm [22]–[26]. Further, to improve the performance of the algorithm, some researchers work on finding more reasonable algorithm parameters [22]–[25] while others focus on increasing the execution speed of the algorithm with the help of hardware [26].

III. ACQUISITION OF EXPERIMENTAL DATA

The parking location information of the Mobikes can be accessed by sequential access to Mobike's APP and its WeChat applet. The grab software ' Packet Capture' was employed to carry out HTTP request and capture packets. This software does not need the root authority to capture data about requests and responses. By analyzing the request data (Figure 4) and response data (Figure 5), the request address of the background services of Mobike can be extracted.

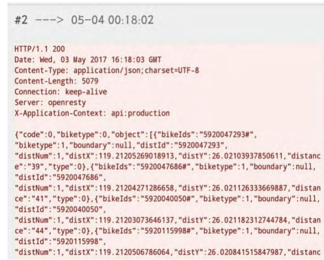


#1 <--- 05-04 00:18:01

POST /mobike-api/rent/nearbyBikesInfo.do HTTP/1.1
charset: utf-8
platform: 4
eption: 1ab88
open_src: list
time: 1493828281604
mobileno: 15880106516
lang: zh
content-type: application/x-www-form-urlencoded
referer: https://servicewechat.com/wx80f809371ae33eda/23/
citycode: 0591
accesstoken: 5f457c3980a6ce86c9a3f0b4f8238cd4
User-Agent: MicroMessenger/6.5.7.1041 NetType/WIFI Language/zh_CN
Content-Length: 96
Host: mwx.mobike.com
Connection: Keep-Alive
Accept-Encoding: gzip

latitude=26.021038&speed=0&errMsg=getLocation%3Aok&longitude=119.21244&ac
curacy=40&citycode=0591

FIGURE 4. The request data.



#2 ---> 05-04 00:18:02

HTTP/1.1 200
Date: Wed, 03 May 2017 16:18:03 GMT
Content-Type: application/json;charset=UTF-8
Content-Length: 5079
Connection: keep-alive
Server: openresty
X-Application-Context: api:production

{"code":0,"biketype":0,"object":[{"bikeIds":"5920047293#",
"biketype":1,"boundary":null,"distId":"5920047293",
"distNum":1,"distX":119.21205269018913,"distY":26.02103937850611,"distanc
e":"39","type":0},{"bikeIds":"5920047686#","biketype":1,"boundary":null,
"distId":"5920047686",
"distNum":1,"distX":119.21204271286658,"distY":26.021126333669887,"distan
ce":"41","type":0},{"bikeIds":"5920040050#","biketype":1,"boundary":null,
"distId":"5920040050",
"distNum":1,"distX":119.21203073646137,"distY":26.021182312744784,"distan
ce":"44","type":0},{"bikeIds":"5920115998#","biketype":1,"boundary":null,
"distId":"5920115998",
"distNum":1,"distX":119.2120506786064,"distY":26.020841515847987,"distanc

FIGURE 5. The response data.

Then, the request addresses of the two background services are obtained. The request address of the WeChat applet is https://mwx.mobike.com/mobike-api/rent/nearbyBikesInfo.do. Its request parameters are longitude and latitude. The request address of the mobile software is https://api.mobike.com/mobike-api/rent/nearbyBikesInfo.do. After obtaining the addresses of the two background services, we further assess their validity and accuracy.

POSTMAN is a powerful Chrome plug-in for web debugging and sending HTTP requests, which can be used to simulate HTTP requests. With its help, the accuracy of the background service address of the mobikes can be verified by comparing with the response data of the network packets. The response data also show that the target data needed in this paper are stored in the object of the reply message in the JSON format.

In this paper, the experimental period of the data collection is one week (from October 9, 2017 to October 15, 2017). The grabbing areas are the five districts (Jin'an, Mawei, Gulou, Taijiang and Cangshan) of Fuzhou city. The record form of each bicycle's information is (grabbing time, bicycle's ID, bicycle's longitude, bicycle's latitude). Considering that each crawler may not be able to get the information about all bicycles, in order to increase the accuracy of our experimental data, the data of seven days are integrated together and then the bicycles' IDs are chosen as the primary key for removing the duplicate bicycles. Finally, it is concluded that the number of bicycles in the five districts of Fuzhou city is 94,310. The specific quantity distribution is shown in the following table (Table 3).

TABLE 3. Number distribution of mobikes in the five districts.

Districts	The number of mobikes
Cangshan	33383
Gulou	18253
Jin'an	24124
Mawei	3182
Taijiang	15368

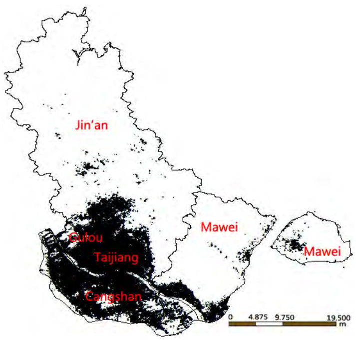


FIGURE 6. The scatter diagram of the distribution of mobikes.

We imported our data into ARGIS software, and then the nuclear density is adopted to analyze them and draw the scatter diagram (Figure 6).

From the figure above, we can see that these bicycles distributed in Gulou district and Taijiang district is far greater than that of others. We argue that the difference is related to the economic development, topography and population distribution of each district.

1) According to the statistics of the South wealth network, in 2017, the population of Gulou district, Taijiang district, Cangshan district, Jin'an district and Mawei district was 826,000, 320,000, 460,000, 510,000 and 156,000, respectively [27].
2) Jin'an district has the largest area and large population. Its main landforms are hills and plains. There are three famous mountains, i.e., Gushan Mountain, Beifeng Mountain and Shoushan Mountain, in the central and northern regions of Jin'an, meaning that this district presents "three high, two low" terrain features. There are also Changle Mountain, Kangshan Mountain, Jinji Mountain and other tourist attractions. These objective factors make the area unsuitable for population concentration, and thus loses its advantages of being a center of Fuzhou city. This also affects the number of shared bicycles put here.
3) Taijiang district is the financial center of Fuzhou city. In 2014, the Fujian provincial government approved the establishment of the Fuzhou-Haixi modern financial center area, and Taijiang district is one of the business districts. Insisting on precise investment promotion is the 'fine tradition' of Taijiang district. Financial institutions of different levels and categories came in a rush, and the agglomeration effect of the financial industry has been prominent. According to a relevant person in charge of Taijiang district, 44 projects have been landed in the Fuzhou-Haixi modern financial center area, with a total investment of 60 billion yuan. A headquarters economy and a building economic agglomeration development platform have taken shape. At present, the tax revenue of Taijiang's financial industry exceeds 20 billion yuan, accounting for 25% of that of Fujian province [28]. The developed financial industry has contributed to the population aggregation to a large extent.
4) The government of Gulou district has been focusing on its industrial structure adjustment. Relying on business building and functional and regional facilities, it attracts all kinds of enterprises, develops the modern service industry by leasing and selling, develops and operating buildings, and cultivates new tax sources and new economic growth points to enhance its dominant position and core competitiveness in Fuzhou city [29]. At present, Gulou district has formed a core area based on the financial industry, with the spatial distribution pattern in the periphery of real estate, wholesale and retail, and the financial industry. Although the geographic area of Gulou district is not very large, it has gathered a larger number of people because the provincial government and a large number of well-known primary, junior and high schools are located there. There is a large number of people gathered in a relatively small areas creating a high-density distribution of sharing bicycles.

From Figure 6, we can also find that Mobikes are evenly distributed in the fringe areas of some districts. Most of these fringe areas are located along the river, close to most of the surrounding roads and convenient transportation. In general, the distribution of Mobikes in the five districts of Fuzhou city can be summarized as follows: their main coverage areas are economic and financial centers, high-density residential centers, flat terrain landscapes and convenient traffic areas. In view of the large difference in Mobikes' coverage in each district, it is necessary to separate the sample data sets of the five districts before selecting service sites. Sequentially, service sites selection is carried out separately in district units.

IV. SERVICE SITES SELECTION AND THE SHORTEST CIRCUIT PLANNING

A. MOBIKES' SERVICE SITES SELECTION BASED ON DENSITY CLUSTERING ALGORITHM

In order to provide more convenient management and maintenance for nearby Mobikes, this paper adopts a density-based clustering algorithm to determine the optimal number and location of service sites in each district.

1) A BRIEF INTRODUCTION OF DENSITY-BASED CLUSTERING ALGORITHM

DBSCAN (Density-Based Spatial Clustering of Application with Noise) is a density clustering method based on high-density link areas. Through a continuous connection of the boundary points with the high-density points in the neighborhood, the clusters are excavated. The size of the cluster depends on the density threshold and the neighborhood size. It uses the parameter $(\varepsilon, MinPts)$ to describe the distribution tightness of neighborhood samples. The parameter ε is the neighborhood distance threshold of the samples, while $MinPts$ stands for the threshold of the number of samples in the neighborhood with a sample distance of ε. Assuming that the sample set is marked as $D = \{X_1, X_2, ..., X_m\}$, the definition of DBSCAN density can be described as following:

1) ε-neighborhood: if $X_j \in D$, then the ε-neighborhood contains a subset of samples in D, where the distance to X_j is not greater than ε. This can be defined as $N_\varepsilon(X_j) = \{X_i \in D | distance(X_i, X_j) \leq \varepsilon\}$.
2) Core object: $\forall(X_j \in D)$, if the number of samples in its ε-neighborhood $N_\varepsilon(X_j)$ is not less than $MinPts$, then X_j can be seen as a core object.
3) Direct density: if $X_i \in N_\varepsilon(X_j)$ and X_j is a core object, then there is a direct density between X_i and X_j.
4) Reachable density: for X_i and X_j, if there is a sample sequence $\langle P_1, P_2, ..., P_T\rangle$ satisfying the following two conditions: 1) $P_1 = X_t, P_T = X_t$; 2) there is a direct density between P_{T+1} and P_T. Then, there is a reachable density between X_i and X_j.
5) Connected density: If there is a core object X that can make the densities of X_i and X_j become reachable, then the densities of X_i and X_j can be seen as connected.

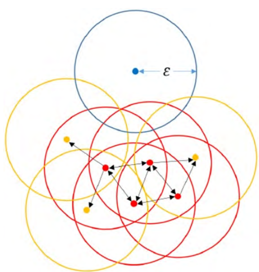


FIGURE 7. A sketch map of DBSCAN.

As shown in the figure above (Figure 7), DBSCAN takes each data point as the center of the circle and draws the circle with ε as the radius. The number of points contained in the circle is taken as the density value of the data point.

Then, an appropriate density threshold ($MinPts$) is set. The point with a density threshold as the center of the circle is marked as X_c. If the density of the data point X_a is less than the density value of X_c, then X_a can be considered as a low-density point. Conversely, if the density of the data point X_b is greater than or equal to the density value of X_c, then X_b can be considered as a high-density point or a core point. When the data overlap between two high-density points, these two points can be connected. If a low-density point is contained in the area of a high-density point, it will be connected to the high-density point and can be seen as a boundary point. Clusters are collectively referred as points that can be linked together, while noise points are low-density points and not belong to any cluster.

2) DETERMINATION THE VALUE OF THE TWO KEY PARAMETERS $(\varepsilon, MinPts)$

DBSCAN Algorithm is very sensitive to the values of the two input parameters $(\varepsilon, MinPts)$. A subtle difference in their values are likely to lead to a large difference in the clustering results. At present, the values of these two parameters are basically determined by experience, which not only leads to a heavy workload for researchers but also makes it is difficult to obtain their accurate values.

For each point p_i in the given data set P, it calculates the distance between p_i and all the other points in P. After sorting the calculated distances from small to large, the ordered set of these distances $Dis_i = \{dis_i1, dis_i2, ..., dis_i(k-1), dis_ik, dis_i(k+1), ..., dis_in\}$ is obtained, where dis_ik is the k-distance of p_i. By the same way, the k-distance of each point in the cluster set can be calculated, sequentially the k-distance set ($E = \{e_1, e_2, ..., e_n\}$) of all points can be obtained.

TABLE 4. The value of n and set partition results for each district.

Districts	The number of original data points	The value of n	The number of partition sets
Cangshan	317447	45	4589
Gulou	111611	30	2924
Jin'an	193317	40	3306
Mawei	58825	25	1605
Taijiang	86399	30	2282

In order to analyze the appearance of the Mobikes in the five districts of Fuzhou city during the experimental period, in this paper, the longitude and latitude of each bicycle are used as the primary keys to remove the duplicate data. Mobikes with the same ID but appearing in different locations are seen as different Mobikes. By displaying the Mobikes with different locations on the map, we can find that many points are distributed intensively. They are clustered into circles and very close to each other. Furthermore, the radius of these circles can be found to be around 0.00023 by calculation. Considering that there may be a condition in which a Mobike is not used but only slightly moves its position, this paper calculates the distance between each individual data point in the sample set and puts the points with distances less than 0.00023 into a set. When the number of elements in the set is greater than n, all the points in the set are taken as one data point. Because the distribution of Mobikes in different districts is inconsistent, in order to make the experimental results more accurate, the values of n in each district have been recorded several times. When the sum of all the data points in the partition sets is closest to the number of the original data points, the value of n in this case will be used as the final value of n. Finally, we can obtain the value of n for each district (Table 4).

The original data sets of each district are divided into groups according to the corresponding values of n, and then the the center points of each partition set is calculated. The center point of partition set ps_i will be used to replace ps_i. By this way, a new sample set can be obtained. DBSCAN will be carried out on the new sample set. After that, to obtain the k-distance set of each district, we calculate the k-distance of all points in the new sample set of each district. Next we take Gulou district as an example to introduce the solution process of the k-distance.

To generate the scatter gram (Figure 8), the increasing natural number sequence is taken as the X axis, while the k-distance of all data points in Gulou district is used as the Y axis to generate a scatter gram (Figure 8). The figure shows that in the interval [2800, 3000), the k-distance rises rapidly and fluctuates greatly, which is not conducive to observing the distance of most data points. Therefore, when calculating the k-distance of Gulou district, we remove the interval [2800, 3000) and retain the interval [0, 2800) (Figure 9).

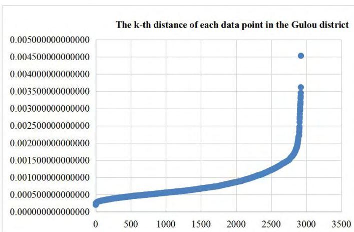


FIGURE 8. The k-distance scatter diagram of each data point in Gulou district.

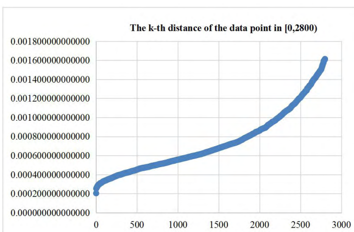


FIGURE 9. The k-distance scatter diagram of the data points in [0, 2800) in Gulou district.

As shown in Figure 9, the k-distances of the data points in Gulou district are mostly located in interval [0.0006, 0.0012]. All the values of k-distances are used as the value of the parameter ε of the DBSCAN algorithm. The number of clusters and low-density points corresponding to each k-distance are calculated and recorded. The appropriate k-distance is selected under the evaluation criterion that the best clustering results can be obtained, which means the least number of points are excluded from the clusters. Meanwhile, the value of *MinPts* corresponding to the best k-distance is also recorded. By this way, we can obtain the values of ε and *MinPts* for each district (Table 5).

TABLE 5. The value of ε and *MinPts* for each district.

Districts	ε	$MinPts$
Cangshan	0.00135146779148	3
Gulou	0.000885013881128	4
Jin'an	0.001269693379658	3
Mawei	0.000824775919085	2
Taijiang	0.000999276086690	8

3) CLUSTERING RESULT AND SERVICE SITES SELECTION FOR EACH DISTRICT

After obtaining the value of k-distance and *MinPts*, the core points of each district can be attained. One point can be viewed as a core point if the number of other points (n) in its $N_\varepsilon(X_j)$ can satisfy the condition of $n \geq MinPts$. Considering that some points may be mistakenly put into the outlier set in this process, we examines the outlier sets and compares them with the sets of core points based on their mapping relationships. By this way, we can reduce the error probability thus obtain more accurate sets of core points.

After obtaining the core point sets, core points in these sets are connected to generate clusters. The clusters contain not only the connected core points but also the points in their $N_\varepsilon(X_j)$s, while outliers are outside the coverage of the clusters. Based on the combination of breadth and depth examinations, one point p is taken from the core point set S, and then all the other points in S connected to p are put into the set C_1. Then, we delete p and the points in C_1 from S, subsequently obtain a new core point set S'. The same operations will be carried on S' until S is empty. The clustering result of Gulou district is shown in Figure 10.

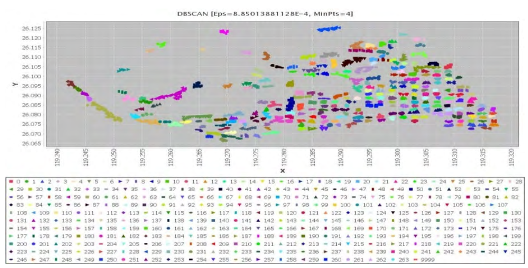


FIGURE 10. The cluster effect diagram of Gulou district.

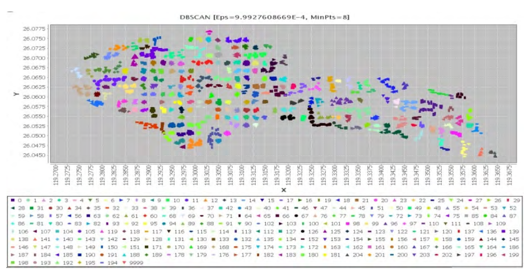


FIGURE 11. The cluster effect diagram of Taijiang district.

After obtaining the final clustering effect of Gulou district, the central point of each cluster can be treated as the location of the ideal service site. The cluster effect diagrams of Taijiang, Cangshan, Jin'an and Mawei district are shown in Figure 11 to Figure 14 respectively. The number of service

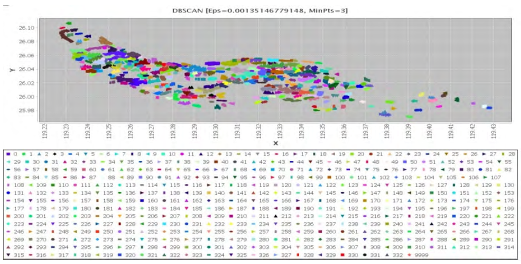


FIGURE 12. The cluster effect diagram of Cangshan district.

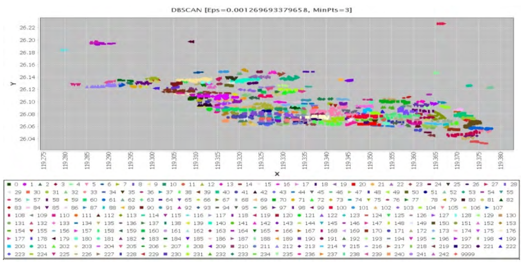


FIGURE 13. The cluster effect diagram of Jin'an district.

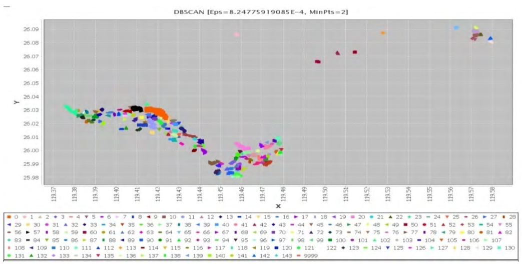


FIGURE 14. The cluster effect diagram of Mawei district.

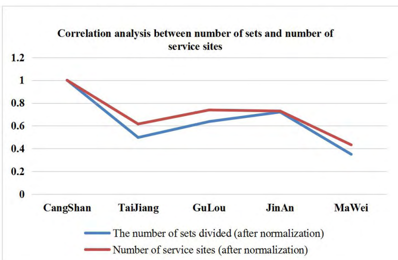


FIGURE 15. Fitting the number of partition sets and the number of service sites in each district.

sites in each district is 246 in Gulou district, 205 in Taijiang district, 333 in Cangshan district, 243 in Jinan district, and 144 in Mawei district.

Normalization of the number of partition sets in Table 4 and the number of service sites in each district is performed to analyze the correlation between the two sets of data. As shown in Figure 15, the correlation between these two data sets is as high as 98.56%, which indicates the validity of the clustering effect and the rationality of the number of service sites.

B. SHORTEST CIRCUIT PLANNING AMONG SERVICE SITES BASED ON THE ANT COLONY ALGORITHM

The service sites can provide management and maintenance for nearby bicycles. How to provide better support and maintenance for these service sites themselves is also the research goal of this paper. After obtaining the locations of service sites in different districts, ant colony algorithm is adopted to find the shortest circuit traversing all of the service sites in a district. Then, we can set the starting points of the shortest circuit to the locations of storage centers. Storage centers in various districts will be responsible for supplying all the service sites in the district or delivering new bicycles to them. Obviously, what we want to study is a typical traveling salesman problem (TSP).

1) A BRIEF INTRODUCTION OF ANT COLONY ALGORITHM

The ant colony system is a bionic optimization algorithm proposed by Yuanyuan and Jing [21]. It can be used to simulate the intelligent behavior of an ant colony in the process of collecting food. They studied the foraging process of ants and found that the behavior of ant colonies was faster and more efficient than that of a single ant. Ant colony releases pheromones as they walk, and by virtue of how pheromones are perceived, they choose to walk along the path of the pheromone concentration, thus ensuring the shortest path to reach the source of food in any environment. The algorithm follows the following rules:

1) Environmental information: environmental information includes obstacles in the environment where ants live, multiple selection paths, nest pheromones and food pheromone. These pheromones are not static but spread at a certain rate in the air.
2) Foraging rule: ants search for food based on pheromones, but not all ants can find the right direction, some ants may move to places with fewer pheromones.
3) Movement rule: when there are no pheromones, ants move according to the inertia of movement direction; otherwise, they will be guided by pheromones.
4) Avoidance rule: when an obstacle occurs in the direction of movement, ants first choose to follow the guidelines of pheromones, otherwise, they will move randomly in other directions.
5) Pheromone emission rule: the pheromones released by ants is inversely proportional to the distance between food and nests. The shorter the distance, the more the pheromones are released.

The relevant parameters in the implementation of the ant colony algorithm are illustrated in Table 6.

2) SHORTEST CIRCUIT PLANNING AMONG SERVICE SITES IN EACH DISTRICT

Combined with the research in this paper, the specific steps of the application of the ant colony algorithm in this paper are as follows:

1) Step 1: Initialize the maximum iteration number and key parameters such as the pheromone factor.

TABLE 6. Key parameters of ant colony algorithm.

Key parameters	Definition
The number of ants	Assuming that the number of service sites is M, the number of ants is m, and the best ratio is 3:2.
Pheromone factor	It describes the relative importance of the amount of pheromone released by ants during the process of searching, and the best value interval is [1, 4].
Heuristic function factor	It is used to describe the relative importance of heuristic information in guiding the search process of ant colony. When the interval is [3,4.5], the algorithm has better performance.
Pheromone volatilization factor	It is used to describe the disappearance level of pheromone, and the comprehensive performance of the algorithm is better when the interval is [0.2 and 0.5].
Pheromone volatilization factor	It is used to describe the total pheromone released during the cycle of an ant cycle. When the value interval is [10, 1000], the algorithm has better comprehensive performance.
Pheromone volatilization factor	Used to control the number of cycles execution, the value cannot be too large or too small: too large, resulting in waste of resources, too small will lead to premature termination of the algorithm.

TABLE 7. Number of service sites and ants used in each district.

Districts	The number of service sites	The number of ants
Cangshan	333	222
Gulou	264	176
Jin'an	243	162
Mawei	144	96
Taijiang	205	136

Do pre-processing, such as converting the coordinate information of service sites to the distance matrix among them.

2) Step 2: Each ant is randomly placed in a different service site. Calculating the next service site that is to be reached until all of the ants have access to all the service sites.
3) Step 3: Calculating the path length of each ant. The optimal solution corresponding to the number of iterations is recorded while the pheromone concentration of the path is updated.
4) Step 4: Determining whether the maximum number of iterations has been reached: if so, then stop the iteration; if not, then return to Step 2.
5) Step 5: Output the result.

Based on the location information of the service sites in the five districts obtained by DBSCAN algorithm, we take the location information of each Mobike as a node on the path. Then, the number of ants (n_{ants}) was determined according to the number of service sites (n_{sites}) in each district. The ratio of n_{sites} to n_{ants} is 3:2 (Table 7).

Research results of [21]–[26] also show that the original ant colony algorithm (OAC) has a shortcoming in solving the TSP problem: with the increase of the number of sites, its running speed become slower and slower and the error between the calculated optimal solution and the real optimal solution become bigger and bigger. To improve the performance of OAC, in this section, we introduce the 2-opt (2-optimization or 2-exchange) [30] and three skills (3-skills). The three skills are described as following:

1) Candidate list: When a ant calculates how to choose the next site, it only calculates a certain number of the nearest sites to current site. If the nearest sits have already been passed, then the probabilistic path selected by ant colony algorithm will be abandoned and replaced by the available nearest cities selected by greedy method.
2) Fixed radius search: When searching for neighborhoods, current site does not need to try to determine whether the connections to all other sites can be exchanged for better results, but only need to try to exchange with several sites closest to it.
3) Don't look bit: If a site can't swap out better results with all its neighbors, then obviously we don't have to search for this site when we are searching for its neighbors.

TABLE 8. Complexity comparison of the three ways.

Algorithms	Algorithm complexity	Explanation
OAC	$O(T * n^3)$	T stands for the number of loop which usually is a large value; n stands for the number of sites.
OAC+(2-opt)	$O(T' * n^3)$	T' also stands for the number of loop, but its value is far less than T; Even if the value of T' can be ignored, the complexity is still $O(n^3)$.
OAC+(2-opt)+(3-skills)	$O(T'' * n^2)$	T'' also stands for the number of loop, but its value is far less than T.

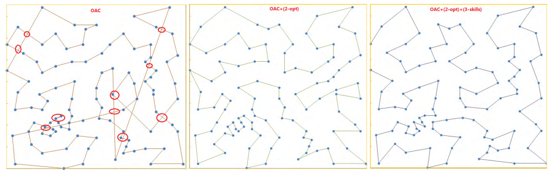


FIGURE 16. The optimal path for 'eil101.tsp'.

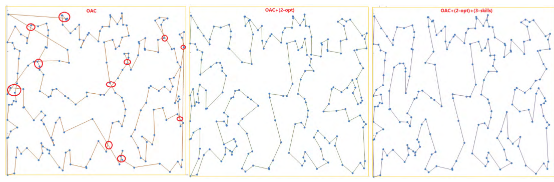


FIGURE 17. The optimal path for 'kroA200.tsp'.

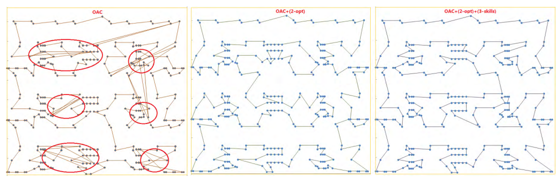


FIGURE 18. The optimal path for 'lin318.tsp'.

From our experiments, we can obtain the complexity of the three ways (OAC, OAC+(2-opt), OAC+(2-opt)+(3-skills)) used for solving the TSP problem (Table 8). Figures 16 to 18 show the difference among the optimal paths calculated by the three ways intuitively.

From the three figures above, we can find that there may be some mistakes (circled in red) in the optimal solution calculated by using OAC alone while OAC + (2-opt) and OAC + (2-opt) + (3-skills) can bring us satisfactory solutions. It should be explained that the differences in longitude and latitude values of the selected service sites may be not obvious, even if they are far away from each other in reality. If we draw the optimal path among them directly, most lines may be very dense, thus making it difficult to observe and compare. So instead, we use data from the three files named 'eil101.tsp', 'kroA200.tsp'and 'lin318.tsp' to make Figures 16 to 18. These files are widely used by

TABLE 9. The best maintenance circuits for service sites in each district.

Districts	Optimal circuits
Cangshan	323, 279, 37, 86, 136, 129, 257, 272, 248, 60, 197, 160, 107, 49, 186, 187, 2, 67, 321, 81, 124, 231, 88, 311, 79, 271, 149, 130, 172, 78, 83, 222, 252, 34, 302, 298, 43, 131, 126, 55, 46, 238, 75, 218, 95, 303, 115, 10, 297, 157, 289, 209, 242, 30, 263, 264, 306, 14, 204, 230, 137, 112, 226, 316, 166, 162, 152, 216, 105, 255, 99, 235, 301, 253, 261, 167, 305, 150, 199, 0, 251, 291, 300, 4, 324, 265, 258, 195, 312, 190, 192, 327, 72, 26, 156, 125, 277, 330, 58, 108, 228, 275, 103, 118, 24, 193, 294, 267, 237, 155, 269, 109, 6, 66, 161, 221, 164, 133, 27, 11, 96, 159, 13, 176, 139, 244, 315, 229, 16, 68, 117, 122, 270, 101, 262, 276, 174, 219, 283, 135, 234, 62, 332, 211, 202, 295, 322, 102, 175, 89, 173, 314, 70, 213, 282, 40, 48, 57, 32, 18, 331, 82, 44, 325, 36, 273, 284, 308, 210, 80, 74, 154, 309, 153, 127, 28, 307, 250, 42, 182, 94, 247, 281, 146, 310, 104, 189, 185, 239, 65, 266, 318, 223, 61, 145, 286, 142, 85, 97, 147, 181, 268, 236, 52, 208, 121, 141, 280, 119, 278, 180, 184, 170, 111, 63, 59, 120, 29, 7, 292, 25, 198, 113, 106, 5, 50, 293, 100, 171, 246, 138, 132, 91, 56, 232, 143, 203, 35, 20, 158, 73, 163, 217, 53, 256, 51, 144, 183, 290, 168, 54, 114, 169, 179, 225, 98, 165, 19, 87, 12, 15, 21, 296, 285, 205, 140, 33, 31, 240, 188, 22, 214, 287, 110, 320, 41, 76, 329, 254, 191, 23, 233, 77, 8, 148, 47, 17, 45, 177, 245, 212, 227, 220, 151, 207, 38, 69, 288, 194, 1, 313, 92, 201, 326, 64, 328, 249, 299, 224, 3, 134, 116, 274, 128, 71, 259, 304, 93, 178, 9, 39, 123, 196, 317, 243, 260, 200, 319, 215, 84, 90, 241, 206, 323
Gulou	154, 133, 105, 93, 240, 95, 99, 65, 207, 178, 182, 118, 79, 242, 57, 114, 232, 88, 153, 164, 146, 41, 211, 121, 91, 142, 139, 52, 261, 210, 38, 251, 73, 172, 239, 176, 129, 12, 58, 236, 2, 198, 15, 62, 4, 227, 177, 257, 111, 218, 152, 237, 163, 21, 254, 233, 263, 134, 36, 174, 22, 120, 216, 18, 60, 29, 124, 33, 102, 219, 213, 64, 103, 225, 141, 256, 167, 171, 179, 212, 0, 40, 27, 147, 14, 55, 241, 217, 169, 51, 28, 214, 229, 247, 168, 243, 110, 160, 32, 17, 191, 109, 48, 20, 37, 107, 196, 127, 10, 181, 173, 45, 230, 161, 157, 189, 151, 19, 208, 156, 112, 253, 228, 223, 108, 235, 194, 71, 82, 209, 155, 90, 116, 35, 131, 193, 130, 56, 259, 78, 3, 148, 224, 200, 136, 25, 59, 72, 115, 158, 75, 248, 138, 85, 185, 203, 238, 206, 258, 106, 165, 23, 9, 16, 13, 34, 125, 100, 245, 122, 92, 128, 143, 26, 63, 188, 135, 76, 184, 97, 183, 190, 132, 47, 43, 119, 61, 67, 104, 31, 24, 84, 74, 5, 39, 81, 70, 42, 197, 87, 262, 145, 199, 53, 126, 244, 98, 246, 30, 250, 231, 234, 8, 192, 144, 186, 96, 86, 175, 215, 252, 222, 83, 249, 204, 50, 221, 150, 170, 44, 137, 159, 80, 220, 66, 1, 94, 117, 255, 140, 149, 49, 54, 187, 195, 89, 11, 46, 101, 69, 180, 260, 7, 113, 205, 166, 201, 162, 77, 6, 226, 202, 123, 68, 154
Jin'an	240, 2, 149, 69, 41, 151, 153, 236, 121, 175, 87, 174, 76, 242, 148, 42, 48, 161, 127, 118, 117, 89, 94, 7, 196, 0, 80, 1, 241, 189, 159, 56, 207, 137, 72, 177, 231, 96, 193, 216, 205, 138, 103, 160, 32, 139, 14, 206, 115, 176, 110, 39, 182, 178, 199, 106, 71, 107, 212, 225, 239, 184, 233, 85, 81, 83, 75, 9, 125, 90, 226, 165, 152, 50, 29, 163, 157, 128, 119, 25, 140, 221, 19, 53, 209, 21, 64, 217, 142, 191, 79, 62, 101, 6, 116, 92, 164, 179, 27, 4, 201, 52, 54, 166, 37, 197, 13, 70, 130, 181, 36, 156, 183, 91, 86, 16, 173, 12, 102, 60, 208, 185, 74, 30, 238, 22, 34, 143, 230, 114, 77, 33, 227, 220, 187, 190, 26, 78, 210, 194, 122, 145, 8, 162, 51, 154, 93, 10, 82, 57, 134, 224, 98, 204, 155, 215, 111, 132, 46, 123, 237, 218, 99, 5, 45, 146, 223, 100, 104, 222, 200, 136, 133, 232, 55, 109, 65, 235, 49, 84, 167, 40, 31, 229, 150, 202, 68, 144, 3, 113, 17, 58, 28, 120, 63, 188, 172, 126, 186, 168, 195, 214, 88, 59, 170, 61, 105, 108, 213, 135, 124, 219, 24, 171, 131, 47, 112, 192, 169, 44, 180, 95, 228, 11, 158, 129, 97, 66, 15, 141, 35, 67, 18, 20, 147, 38, 23, 211, 234, 198, 43, 203, 73, 240
Mawei	135, 99, 130, 60, 42, 50, 53, 5, 129, 142, 22, 10, 18, 48, 65, 113, 1, 2, 71, 0, 102, 57, 75, 9, 31, 127, 106, 112, 78, 79, 46, 33, 115, 45, 94, 98, 38, 122, 61, 23, 47, 84, 81, 64, 73, 100, 43, 27, 140, 103, 14, 132, 137, 7, 111, 141, 109, 80, 39, 101, 120, 70, 11, 119, 8, 123, 30, 85, 16, 121, 15, 74, 126, 40, 125, 24, 13, 118, 12, 104, 69, 52, 59, 25, 86, 34, 54, 63, 56, 21, 89, 124, 26, 138, 114, 67, 77, 117, 29, 72, 97, 32, 110, 90, 96, 107, 131, 143, 87, 58, 62, 19, 3, 91, 36, 6, 83, 95, 20, 37, 55, 93, 116, 28, 66, 44, 51, 35, 88, 49, 133, 82, 108, 4, 136, 17, 41, 139, 68, 76, 105, 128, 134, 92, 135
Taijiang	125, 200, 94, 105, 129, 48, 166, 5, 11, 99, 189, 178, 183, 102, 194, 203, 112, 198, 88, 156, 93, 109, 168, 23, 173, 165, 84, 147, 174, 27, 118, 179, 177, 185, 16, 124, 90, 10, 131, 108, 103, 138, 0, 80, 199, 36, 114, 140, 127, 15, 161, 74, 190, 188, 196, 17, 193, 66, 146, 133, 89, 162, 197, 56, 73, 97, 154, 117, 50, 176, 96, 1, 111, 77, 3, 143, 81, 153, 18, 167, 37, 14, 113, 78, 110, 64, 184, 157, 65, 134, 42, 186, 204, 141, 6, 25, 68, 2, 98, 40, 76, 106, 57, 202, 152, 53, 28, 59, 7, 26, 120, 47, 46, 22, 61, 67, 49, 126, 83, 52, 33, 87, 8, 20, 43, 51, 24, 180, 62, 163, 191, 44, 160, 169, 31, 92, 128, 104, 86, 55, 164, 30, 29, 58, 195, 132, 21, 115, 41, 158, 34, 70, 13, 45, 39, 142, 85, 151, 148, 60, 19, 135, 137, 122, 32, 82, 192, 72, 100, 187, 95, 172, 181, 38, 54, 145, 130, 159, 63, 12, 107, 9, 91, 182, 175, 171, 75, 4, 35, 119, 149, 123, 144, 136, 71, 69, 170, 155, 79, 201, 121, 101, 150, 139, 116, 125

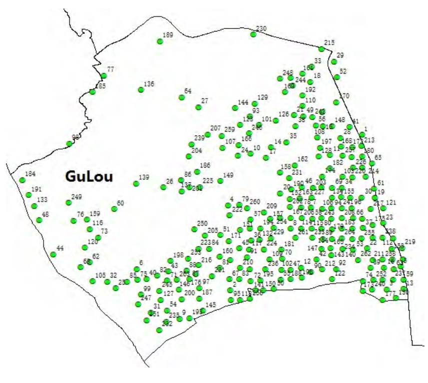


FIGURE 19. The distribution diagram of numbered service sites in Gulou district.

the researchers who study the TSP problem. The number of sites in these files are 101, 200, 318 respectively. They are approximate to the number of selected service sites in each district.

It is important to note that although OAC also gives the solutions to 'kroA200.tsp'and 'lin318.tsp', the time it costs for calculating is unbearable, especially to 'lin318.tsp'. So, taking into overall consideration the accuracy of solution and complexity, there is no doubt that we recommend the third way (OAC + (2-opt) + (3-skills)). Finally, we can

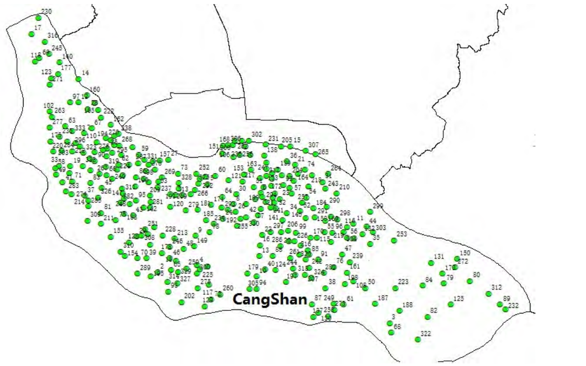


FIGURE 20. The distribution diagram of numbered service sites in Cangshan district.

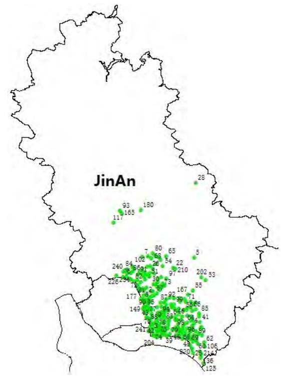


FIGURE 21. The distribution diagram of numbered service sites in Jin'an district.

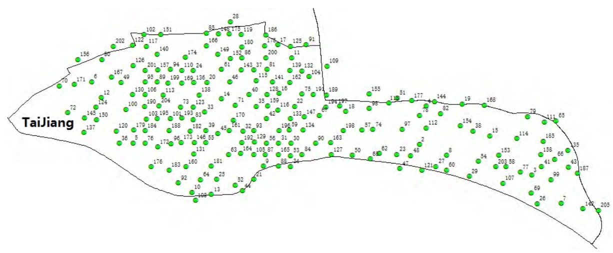


FIGURE 22. The distribution diagram of numbered service sites in Taijiang district.

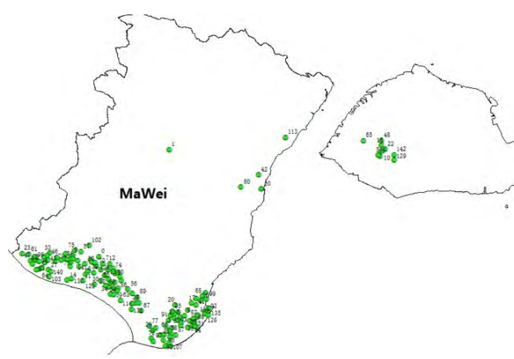


FIGURE 23. The distribution diagram of numbered service sites in Mawei district.

obtain the best maintenance circuits for each regional service site (Table 9). The locations of the numbered service sites in Table 9 are shown in Figures 19 to 23 respectively.

V. CONCLUSION

Bicycle-sharing brings convenience to people's travel but also causes some problems because of various subjective and objective reasons. In this paper, we use the network crawler to collect the location data of the Mobikes' parking spots in the five districts of Fuzhou city. Then, taking these data as the core experimental data, the density-based clustering algorithm is adopted for selecting the optimal service sites for Mobikes in different districts. After finishing this selection work, an improved ant colony algorithm is adopted for planning the shortest circuits passing through all of the service sites in each district. The purpose of doing this is to provide better support for those service sites themselves, such as the provision of parts and technology. The results of our study mainly include:

1) There are obvious differences in the distribution of Mobikes in the five districts of Fuzhou city. The distribution densities in Gulou and Taijiang districts are much higher than that of the other three districts. Mobikes in Fuzhou city mainly cover the regional economic and financial center, the high-density residential center of the population, and the flat terrain and convenient travel areas.
2) According to the clustering results of the location information on the motorbike parking spots in each district, we can find that the number of clusters of Mobikes' centralization parking spots in Cangshan district is the most, followed by Jin'an district. The number of service sites identified in each district is directly proportional to the number of their clusters, and the correlation is as high as 98.56%.
3) Adopting the original ant colony algorithm alone to solve the TSP may not obtain a good solution but cost long running time. Our experimental results prove that introducing 2-opt and 3-skills is efficient for improving the performance of OAC.

To maximize the value of our experimental data, the follow-up work of this paper will be concentrated on the following aspects: 1) Studying and analyzing various clustering algorithms and the TSP problem further. 2) Collecting other kinds of data and making correlation analysis of them with the experimental data of this paper to solve or explain some specific problems and phenomena.

ACKNOWLEDGMENT

The authors would like to thank Professor Jinbo Xiong and Professor Jie Li for their valuable comments. They would also like to thank Professor Xuan Li for the English language revision, which greatly improved the original manuscript.

REFERENCES

[1] *The 41th Statistical Reports on the Internet Development in China*, CNNIC, Beijing, China, 2018.
[2] A. Tian. (Feb. 2018). *The Total Number of Shared Bicycles has Reached 23 Million and its Registered Users is 400 Million*. [Online]. Available: http://finance.eastmoney.com/news/1355,20180208832343045.html
[3] Surging News. (Nov. 2017). *Shanghai has Cleaned up 516 Thousand Bicycles, and its Number now is 1 Million 150 Thousand*. [Online]. Available: http://news.163.com/17/1106/11/D2IB8BT900018AOR.html
[4] X. Lv and L. Zhao. (Mar. 2017). *Dongguan: Amount of Delivery of the Shared Bicycle is About 775 Thousand and the Management Specification for Them is Being set up*. [Online]. Available: http://news.timedg.com/2018-03/17/20649658.shtml
[5] T. Zhang, R. Ramakrishnan, and M. Livny, "BIRCH: An efficient data clustering method for very large databases," *ACM SIGMOD Rec.*, vol. 25, no. 2, pp. 103–114, 1996.
[6] G. Hamerly and C. Elkan, "Alternatives to the k-means algorithm that find better clusterings," presented at the 11th Int. Conf. Inf. Knowl. Manage., McLean, VA, USA, Nov. 2002, pp. 600–607.
[7] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise," presented at the 2nd Int. Conf. Knowl. Discovery Data Mining, Portland, OR, USA, Aug. 1996, pp. 226–231.
[8] W. Wang, J. Yang, and R. Muntz, "STING: A statistical information grid approach to spatial data mining," presented at the 23rd Int. Conf. Very Large Data Bases, San Francisco, CA, USA, Aug. 1997, pp. 186–195.
[9] H. Ichihashi, K. Miyagishi, and K. Honda, "Fuzzy c-means clustering with regularization by K-L information," presented at the 10th IEEE Int. Conf. Fuzzy Syst., Melbourne, VIC, Australia, Dec. 2001.
[10] D. Arlia and M. Coppola, "Experiments in parallel clustering with DBSCAN," presented at the 7th Int. Eur. Conf. Parallel Process., Berlin, Germany, Aug. 2001, pp. 326–331.
[11] S. Liu, Z.-T. Dou, F. Li, and Y.-L. Huang, "A new ant colony clustering algorithm based on DBSCAN," presented at the Int. Conf. Mach. Learn. Cybern., Shanghai, China, Aug. 2004, pp. 1491–1496.
[12] T. Ali, S. Asghar, and N. A. Sajid, "Critical analysis of DBSCAN variations," presented at the Int. Conf. Inf. Emerg. Technol., Karachi, Pakistan, Jun. 2010, pp. 1–6.
[13] A. Smiti and Z. Eloudi, "Soft DBSCAN: Improving DBSCAN clustering method using fuzzy set theory," presented at 6th Int. Conf. Hum. Syst. Interact. (HSI), Sopot, Poland, Jun. 2013, pp. 380–385.
[14] K. Khan, S. U. Rehman, K. Aziz, S. Fong, and S. Sarasvady, "DBSCAN: Past, present and future," presented at the 5th Int. Conf. Appl. Digit. Inf. Web Technol. (ICADIWT), Bangalore, India, Feb. 2014, pp. 232–238.
[15] S. Shamshirband, A. Amini, N. B. Anuar, M. L. M. Kiah, Y. W. Teh, and S. Furnell, "D-FICCA: A density-based fuzzy imperialist competitive clustering algorithm for intrusion detection in wireless sensor networks," *Measurement*, vol. 55, pp. 212–226, Sep. 2014.
[16] I. Cordova and T.-S. Moh, "DBSCAN on resilient distributed datasets," presented at the Int. Conf. High Perform. Comput. Simulation (HPCS), Amsterdam, The Netherlands, Jul. 2015, pp. 531–540.
[17] R. Bellman, "Dynamic programming," *Science*, vol. 153, no. 3731, pp. 34–37, 1966, doi: 10.1126/science.153.3731.34.
[18] R. M. Karp and Y. Zhang, "Randomized parallel algorithms for backtrack search and branch-and-bound computation," *J. ACM*, vol. 40, no. 3, pp. 765–789, 1993, doi: 10.1145/174130.174145.
[19] J. H. Holland, "Erratum: Genetic algorithms and the optimal allocation of trials," *SIAM J. Comput.*, vol. 3, no. 4, p. 326, 1973.
[20] M. Dorigo, V. Maniezzo, and A. Colorni, "Ant system: Optimization by a colony of cooperating agents," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 26, no. 1, pp. 29–41, Feb. 1996.
[21] L. Yuanyuan and Z. Jing, "An application of ant colony optimization algorithm in TSP," presented at the 5th Int. Conf. Intell. Netw. Intell. Syst., Washington, DC, USA, Nov. 2012, pp. 61–64.
[22] Y. Li and S. Gong, "Dynamic ant colony optimisation for TSP," *Int. J. Adv. Manuf. Technol.*, vol. 22, nos. 7–8, pp. 528–533, 2003, doi: 10.1007/s00170-002-1478-9.
[23] J. Zhang, H. Liu, S. Tong, and L. Wang, "The improvement of ant colony algorithm and its application to TSP problem," presented at the 5th Int. Conf. Conf. Wireless Commun., Netw. Mobile Comput., Beijing, China, Sep. 2009, pp. 1–4.
[24] S. K. Sahana and A. Jain, "High performance ant colony optimizer (HPACO) for travelling salesman problem (TSP)," presented at the Int. Conf. Swarm Intell., Hefei, China, Oct. 2014, pp. 165–172.
[25] H. Shi, L. Zheng, and G. Liu, "Research on TSP based on ant colony algorithm," presented at the IEEE Int. Conf. Inf. Automat., Lijiang, China, Aug. 2015, pp. 2048–2051.

[26] J. Wang, J. Dong, and C. Zhang, "Implementation of ant colony algorithm based on GPU," presented at the 6th Int. Conf. Comput. Graph., Imag. Vis., Washington, DC, USA, Aug. 2009, pp. 50–53.
[27] The South Wealth Network. (Dec. 2017). *Population Number List of Fuzhou Districts in 2017*. [Online]. Available: http://www.southmoney.com/paihangbang/201710/1692686.html
[28] Y. Lin and Y. Lin. (Nov. 2016). *Taijiang District of Fuzhou: Accelerate the Construction of the Core Business Area of Commerce in Provincial Capital Cities*. [Online]. Available: http://www.sohu.com/a/119950765_268708
[29] Z. Lin, S. Wei, T. Ning, J. Chen, and Z. Yu, "Industries spatial distribution feature of building economy in gulou district of Fuzhou," *J. Guizhou Normal Univ., Natural Sci.*, vol. 35, no. 1, pp. 30–36, 2017.
[30] G. A. Croes, "A method for solving traveling-salesman problems," *Oper. Res.*, vol. 6, no. 6, pp. 791–812, 1958.

BIAO JIN received the M.S. degree in software engineering from the Faculty of Software, Fujian Normal University, China, in 2011, where he is currently pursuing the Ph.D. degree. He is currently a Lecturer with Fujian Normal University. His research interests include computer networks, geographic information, and data mining.

MEI-HUANG HONG received the bachelor's degree from the College of Mathematics and Informatics, Fujian Normal University. She is currently with Fujian Newland Software Engineering Co., Ltd. She is mainly involved in data mining.

XUE LIN received the bachelor's degree from the College of Geographic and Environmental Sciences, Fujian Normal University, where she is currently pursuing the M.S. degree with the College of Geographical Science. Her research interests include geographic information.

JIN-MING SHA received the Ph.D. degree from the Agricultural Remote Sensing and Information Technology Application Research Institute, Zhejiang University, Zhejiang, China, in 2000. He is currently a Professor and the Ph.D. Supervisor with the College of Geographical Sciences, Fujian Normal University. His research interests include remote sensing application and geographic information system.