Journal of Computers Vol. 31 No. 2, 2020, pp. 227-240
doi:10.3966/199115992020043102019

Analysis and Exploration of Open Source Data in Traffic Network Based on Shared Bicycle Arrangement Model

Xia Hu[1], Jingjing Cheng[1*], Wenjie Zhong[1]

[1] Department of Management, China University of Mining and Technology, Xuzhou 221116, China
Chengjj@cumt.edu.cn

Received 26 May 2019; Revised 29 July 2019; Accepted 6 January 2020

Abstract. The purpose of this study is to study the layout of shared bikes in the traffic network. In this study, according to the open source data provided by AMAP and the specific situation of Beijing, Canopy algorithm is used to analyze the layout and establish a reasonable layout model. The number and location of the required points are recorded and filtered as the k value of K-means algorithm. Then, the shared bicycles are arranged by clustering results to find the optimal layout strategy. The results show that the cluster analysis of demand points using Canopy algorithm and K-means algorithm can find that there is a large demand in residence, subway station and office area. Therefore, in the future layout of shared bicycles, it is necessary to focus on these areas. From this, it can be seen that using open source data to study the shared bicycle layout model in traffic network is of great significance to alleviate traffic congestion in urban areas.

Keywords: canopy clustering, data analysis, hadoop, layout model, open source data, shared bicycle

1 Introduction

With the continuous development of China's social economy, the level of urbanization is getting higher and higher, and people' s economic level has also improved. Private cars have become the first choice for most people and families to travel. At the same time, the increasing number of private cars has also brought many problems to urban public transport, such as traffic congestion, air pollution and so on. As a result, governments in many places have developed and established public bicycle systems, alleviating traffic pressure and air pollution to a certain extent, and contributing to solving the "last kilometer" problem of public transport. But many places still have the situation of "no place to borrow cars" and "no place to return cars". With the emergence of shared bicycles, the development of public transport has been boosted. It has also become a shared and intelligent public transport development concept which combines with traffic rules, urban development and comprehensive management, and has realized the comprehensive coverage of urban transport network. Shared bicycle also attracts a lot of users because of its advantages of low carbon and environmental protection, and bicycle can be retrieved from and return to an unfixed location. However, after the use of bicycles, there will be a lot of accumulation in some places, and the company cannot schedule and manage in time, which makes the shared bicycle not available in other places. Shared bicycle parking disorderly and a large number of accumulation even cause certain pressure on urban traffic, seriously affecting the travel of other means of transport [1].

At present, sharing bicycle has become a new type of green and environment-friendly sharing economy, and the government is paying more and more attention to the layout and development of shared bicycle in the city. Therefore, facing the situation of shared bicycles parking in disorder, various renovation activities have also been carried out. Beginning in June 2019, on the basis of carrying out special renovation action on Internet bicycle rental, Beijing has incorporated all the important roads in the city into the implementation scope of "Rulan Settlement" and regulated the use of vehicles to a

* Corresponding Author

certain extent [2]. For sharing bicycles, a lot of data can be obtained from AMAP, Baidu Map and related bicycle sharing websites, which is convenient for us to analyze and study the data. Considering that shared bicycle will be put into use in many cities and a lot of data can be obtained, K-Means is used with Canopy algorithm in clustering analysis to forecast the demand of shared bicycle, and then analyses and studies the layout model of shared bicycle to find the optimal layout model, so as to alleviate traffic congestion, reduce environmental pollution, and finally achieve the reasonable layout of traffic network and efficient travel of urban traffic. At the same time, it is also helpful for the management of the city, as well as for the co-creation of green transportation and civilized cities [3].

Using open source data to analyze the layout model of shared bicycles in urban traffic network is not only helpful to alleviate urban traffic, but also significant to urban environment. Therefore, increasing the number of shared bicycles around the bus station and the subway station makes the transfer between the bus and the subway more convenient, which helps to improve the travel efficiency of the entire network and meet the travel needs of the residents. It is helpful to improve the robustness of the network, optimize the urban public transportation network system, and promote the healthy and orderly development of the urban transportation system by analyzing different networks with complex network theory, exploring and protecting important nodes and connected edges in the network.

2 Literature Review

2.1 Intelligent Transportation Network

Traffic network is closely related to people's lives. Most of the research on traffic network by domestic and foreign scholars focuses on the empirical analysis of network and the establishment of network model. In the aspect of urban road network, Griswold et al. divided streets according to the postal codes of different regions and construct the street topology network. The empirical results show that the network has the characteristics of small world. Porta analyzes the road traffic networks of six cities and found that these networks all have small world characteristics [4]. Cen et al. contrasted and analyzed the self-organized urban road traffic network and the planned urban road traffic network. According to the analysis of relevant indicators, it is concluded that the self-organized city has the same scale-free characteristics as the non-spatial network [5]. Amini et al. analyzed the road network topological characteristics of more than twenty big cities in Germany. The results show that traffic flow mainly concentrates on a few roads in the road network, and presents the power law distribution characteristics, which shows that the road network has the characteristics of grading [6]. Andrés et al. described the parameters of urban environmental network, central structure and demand mode, find the best spatial layout of public transport lines, and can clearly find out the relationship between urban environment and user characteristics, so as to solve the layout problem [7]. Liu analyzed the application of intelligent computer technology in urban traffic signal system. The results show that the application of agent technology in traffic signal control system can improve intelligent traffic management, provide important technical support for alleviating urban traffic pressure, and run more comprehensively, safely and reliably in future applications [8].

2.2 Shared Bike

Shared bicycle system is first proposed by Amsterdam. The research and development of shared bicycle in Europe has always been in the leading position. Shared bicycle starts late in China, but its development is very rapid. The research on shared bicycle mainly focuses on network transfer and location planning. Han et al. took shared bicycle as an example to discuss the sharing economy in China. The results show that the business model of shared economy has accelerated the economic development of China, and the management of shared bicycle has also provided some suggestions for future shared products [9]. Cao took Lanzhou as an example to discuss the impact of shared bicycle on urban traffic trip. The results show that the age, occupation and educational background of urban residents will affect their choice of bicycles and a series of parking and returning behaviors to a certain extent. Moreover, their compliance with traffic rules has a direct impact on urban traffic. Finally, in view of this situation, some improvement suggestions and measures are put forward [10]. Yasee et al. analyzed the relevant government policies and the actual transfer situation, elaborate how to transfer effectively, and through the comparative analysis of the use of several countries, analyze the important factors affecting the development of transfer. Based on the heat of location and parking facilities, Horner puts forward relevant layout schemes. The results show that the functions of different transfer points are different [11]. Liu et al. found many problems in the shared bicycle system by studying the shared bicycle system in Beijing, and gives some suggestions [12]. Jarosh et al. summarized and analyzed the shared bicycle system in Hangzhou, which is the first one to use shared bicycle, and put forward some suggestions, which can be used as reference for the establishment and use of shared bicycle system in other cities [13].

Tong et al. used cluster analysis as the research method to make an empirical analysis of the relevant indicators data of 11 major cities in Hebei Province, and give some positioning and location suggestions, which have certain reference significance for the planning and decision-making of logistics parks [14]. Through the analysis of the inter-regional layout structure and connection strength of Hebei Province, and clustering analysis of the network structure characteristics, Zhou et al. found that these regions to a certain extent have "Matthew effect", and the spatial layout needs to be optimized urgently, so as to play a good synergistic role for the development of Beijing-Tianjin-Hebei region [15].

2.3 Sharing Economy

Clarke et al. found that with the continuous development of the sharing economy, various emerging Internet companies are attracted by the public's attention and funds. In recent years, the sharing bicycle industry has become an important part of the sharing economy [16]. From the second half of 2016 to the first half of 2017, the shared bicycle industry has completed multiple rounds of financing and is booming. In this research, the business model of shared bicycles is studied. Through the analysis of the operation mode, profit model and marketing model of shared bicycle, it is found that its profit mode is relatively single, with unreasonable regional distribution and lack of competitive marketing and other problems.

Research by Zhang et al. show that with the expansion of intelligent transportation system (ITS) in smart cities, as a new green public transport mode, shared bicycles are rapidly developing and changing the travel habits of citizens all over the world, especially in China [17]. The purpose of this study is to provide an inclusive review and survey of shared bikes, in addition to its benefits, history, and brand. Moreover, everybody, it is the first time that it has come up with the concept of the shared bicycle Internet (IoSB) to find a viable solution to the technical problem of sharing bicycles. The possible architecture of IoSB, most of the key Internet of Things (IoT) technologies, and their ability to integrate and apply to different parts of IoSB are proposed. At the same time, some of the challenges and obstacles to the implementation of IoSB are thoroughly explained. In terms of recommendations for overcoming these barriers, the potential aspects and applications of IoSB for future technology development in smart cities provide another valuable discussion for this research.

To sum up, the shared economy has become a part of the economy in China, and the shared bicycle has become an important tool for the last kilometer of urban traffic. Previous literatures mostly focus on the study of shared economy and urban traffic network, but seldom on the situation of shared bicycle in urban traffic. Therefore, in this study, open source data is used to analyze the demand for shared bicycles in the traffic network, and then build the layout model of shared bicycles. The aim is to better manage urban traffic.

3 Research Model

3.1 Shared Bicycle in Traffic Network

Traffic network refers to the whole traffic network composed of various transport networks and postal power grids, also known as transport network [18]. Facility network, route network, organization network and demand network constitute the urban traffic network system. Interwoven networks form the important representation of the spatial interaction of human social and economic activities and even the urban architecture. Among them, traffic nodes form a network of facilities, traffic lines form a route network, and the combination of nodes and lines forms a traffic organization network [19].

In recent years, with the improvement of urban level, the number of motor vehicles has increased sharply. Traffic congestion and environmental pollution have emerged in many places. Building a sustainable urban transport system has become a major basic economic and social problem. Sustainable traffic construction should not only describe the mechanism of congestion formation from the micro perspective, but also dynamically analyze the whole traffic network from the macro perspective, and study the interaction between the large system and the traffic network [20].

Shared bicycle is one of the traditional modes of shared transportation. In view of the lack of a unified definition of shared bicycle in academic circles, it is considered that shared transportation is a form of carpooling and network contract vehicles, and it is also considered as a new mode in modern transportation network [21]. In the 1940s, shared transportation had already sprouted in Europe. At that time, because cars were still new things and their prices were relatively expensive, people jointly purchased and used them together, which saved costs to a large extent. With the popularity of automobiles, private cars became the preferred means of transportation for many families, which also made traffic congestion and increased the pressure of the environment. Therefore, fully aware of these points, the United States took the lead in promulgating decrees to actively guide the development of urban traffic to mass rail transit, followed by the car-sharing, high-load-bearing car system. Now, it has become the rudiment of shared transportation [22].

What's more, with the improvement of people's living standards, people's awareness of environmental protection is also growing. Therefore, low-carbon travel has become a choice for people to go out. Shared bicycle developed earlier abroad, and it first appeared in Hangzhou in 2016. A bicycle boom is set off by Mobike Bicycle Company, so many enterprises begin to enter the shared bicycle industry [23]. After several years of development, by January 2019, China's shared bicycle user area has also undergone some changes, as shown in Fig. 1 and Fig. 2.

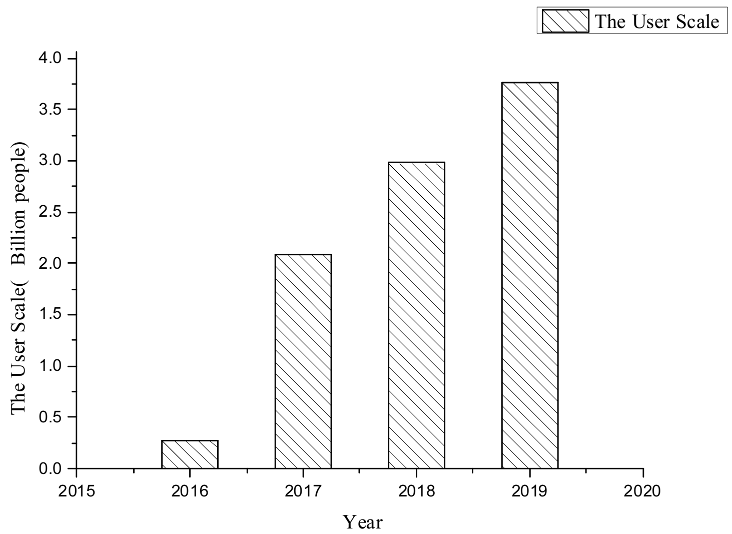

Fig. 1. Trends in the size of shared bicycle users in China from 2016 to 2019

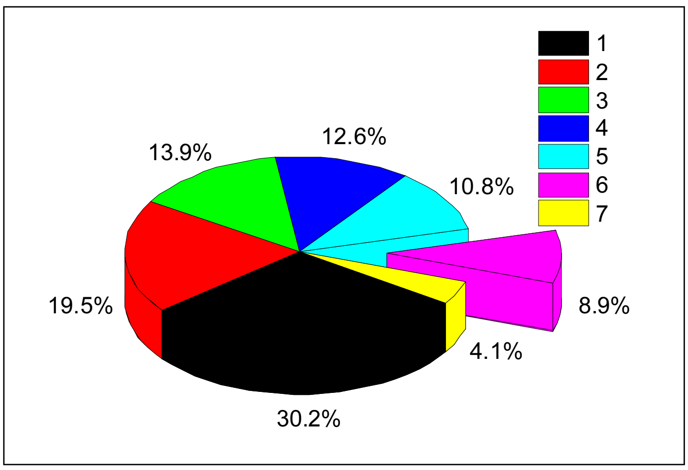

Fig. 2. Regional distribution of shared bicycle users in China in the first quarter of 2019

Among them, 1 represents East China, 2 represents Central China, 3 represents North China, 4 represents South China, 5 represents Southwest China, 6 represents Northwest China and 7 represents Northeast China.

Especially in 2018, after the re-shuffling of shared bicycles, the market share of Harrow bicycles has surpassed other brands and increased to 50%.

The emergence of shared bicycles has provided new ideas for the construction of multi-level urban transportation system. There are several main characteristics. Firstly, cycling is mainly short-term and short-distance. According to the statistical data of Mobike bicycles, most of the bicycles ride is 10 minutes in winter, 12 minutes in spring. Most of the bicycle ride distance is less than 2 km. Secondly, sharing bicycles has remarkable connecting function. Thirdly, sharing bicycles help alleviate congestion. Fourthly, the main roads bear a larger amount of cycling. Fifthly, the cohesion of early peak is significant, and the expansion of late peak is significant. Therefore, the high density of bicycle-sharing stations complements the shortcomings of Metro and bus very well, and becomes the most ideal mode of transportation for people to realize the “last kilometer” traffic from residential to bus, metro stations and different means of transportation [24].

The rapid development of shared bicycle not only has enough quantity in the first-tier cities, but also saturates in the second-tier cities. At the same time, many shared bicycle enterprises also have some problems such as confusion in management and disorderly parking, which put the shared bicycle enterprises in a difficult situation of development. At present, the more feasible scheme is to adopt electronic fencing to specify parking points. If the later planning is not reasonable, the situation of sharing bicycles will be the same as that of traditional public bicycles, which will eventually cause a large number of users to abandon the use [25].

3.2 Data Analysis Tool

Data analysis refers to the process of analyzing a large amount of data collected by appropriate statistical methods, extracting effective information and forming conclusions, and studying and summarizing the data in detail, which can help people make decisions so as to take action. There are many sources of data, such as search engine spider crawl data, website IP/PV and other basic data, HTTP response time data, website traffic source data and so on. With the development of “Internet +”, big data are more and more widely used in people’s lives. When studying the layout of shared bicycles in transportation networks, the open source data of AMAP can be used and bicycle related websites can be shared to analyze and facilitate the accurate grasp of the popular use areas of shared bicycles, and timely increase or decrease the amount of shared bicycle, and realize the effective sharing bicycle scheduling [26].

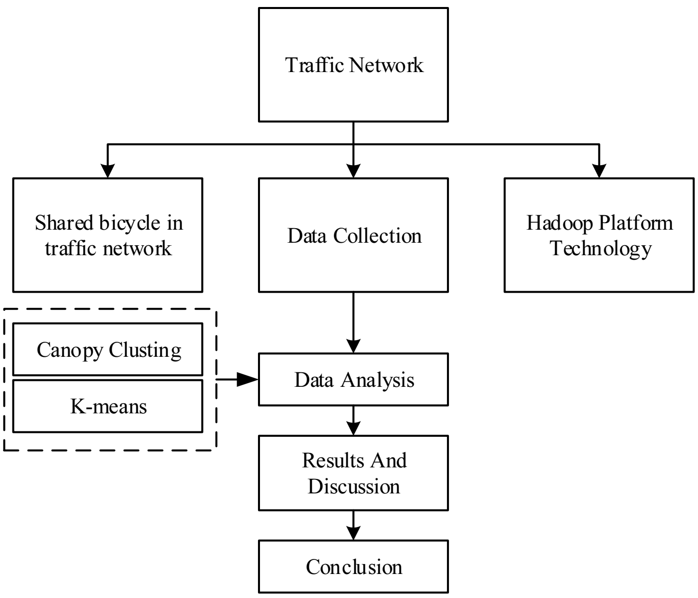

Fig. 3. System process diagram

Clustering analysis means that similar objects are classified into different groups or more subsets by static classification method, so that members in the same subset have some similar attributes, and then these data are analyzed. Canopy algorithm is a clustering algorithm for “big data”. It is suitable for the situation of large number of classes, large number of samples and high dimension of counterfeits. The flow chart of Canopy algorithm is shown in Fig. 4 [27]:

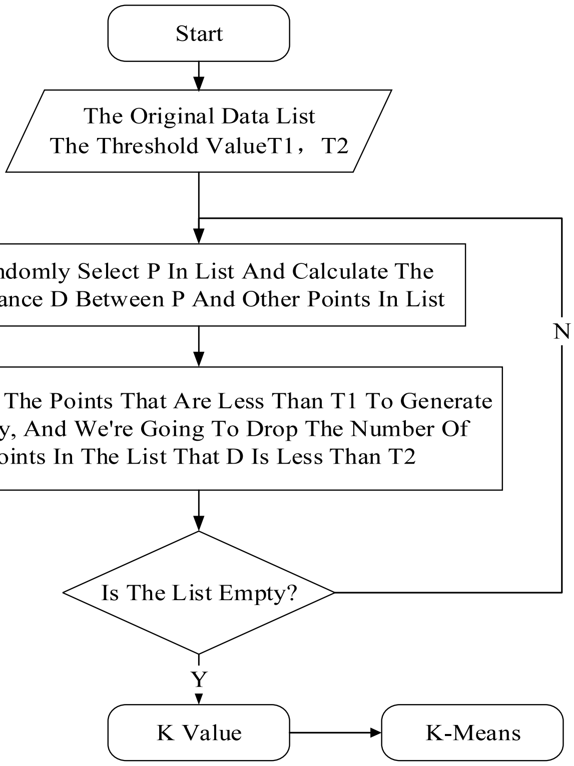

Fig. 4. Flow chart of Canopy algorithm

According to Fig. 4, requirements can be analyzed. The number and location of demand points are recorded and filtered as K value of K-means algorithm, and then shared bicycles are arranged by clustering results.

The calculation process of Canopy algorithm only considers overlapping data vectors, so it can simplify the calculation process, which can not only greatly improve the efficiency, but also save more events. The clustering diagram of Canopy algorithm is shown in Fig. 5.

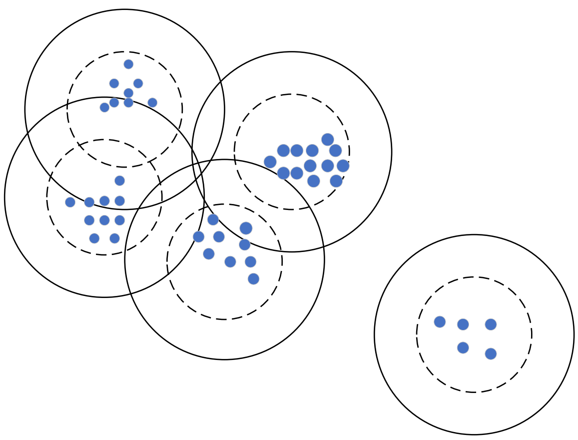

Fig. 5. Canopy clustering graph

K-means algorithm is a clustering algorithm based on unsupervised learning, which is widely used in many fields. In the practical application process, the similarity between different objects is judged by calculating Euclidean distance. The similar objects are classified into one class according to the similarity situation. The key step in K-means is to set up the number of clusters K and initial cluster centers in advance for users, and calculate the Euclidean distance between data vectors and cluster centers in the process of iteration calculation, as shown in Equation 1, so as to obtain new cluster centers, as shown in Equation 2.

$$d(x,y)=\sqrt{(x_1-y_1)^2+(x_2-y_2)^2+\cdots+(x_n-y_n)^2}=\sqrt{\sum_{i=1}^{n}(x_i-y_i)^2} \quad \textbf{(1)}$$

$$J_c=\sum_{j=1}^{k}\sum_{i=1}^{n_j}\left|X_i^j-c_j\right| \quad \textbf{(2)}$$

3.3 Hadoop Platform Technology

Hadoop is a distributed system infrastructure, which is first proposed by Doug cutting. It can provide users with a transparent and detailed distribution of the underlying system. Hadoop has a good cross-platform feature, making full use of the power of the cluster for high-speed operation and storage. Hadoop implements a distributed file system (HDFS) called HDFS. On the one hand, it has the characteristics of high fault tolerance and is designed to be deployed on low-cost hardware. On the other hand, it can provide high throughput to access application data, which is suitable for applications with large data sets. The core design of Hadoop framework is HDFS and MapReduce, in which HDFS provides storage for massive data and is the open source implementation of Google File System (GFS), while MapReduce provides computation for massive data and is the open source implementation of Google MapReduce [28-31].

HDFS adopts a Manager-Worker structure model. The HDFS cluster consists of a unique name node and a large number of data nodes, that is, managers and workers. The manager can map the data block to the worker, who can process the HDFS client read/write request accordingly, and add, delete and duplicate the data block according to the information indicated by the former [32].

MapReduce has two key components, which are Map and Reduce. The Map function maps a set of key-value pairs into a new set of key-value pairs, and the Reduce function aggregates these data to form a scalar. By embedding MapReduce programming into Hadoop clusters, large data sets can be processed in parallel. In the process of MapReduce input, large data sets need to be partitioned to get a number of small data blocks. Map process processes all small data blocks from different nodes in parallel, while completing effective sorting. Reduce process requires specification of data to output the relevant processing results. Read and write content is usually saved to HDFS. YARN not only manages the underlying resources, but also schedules the platform application reasonably. If the application starts up, it must be restarted again. In MapReduce framework, data storage and computing nodes are basically the same, which can reduce the need for cluster broadband, and read the required data directly from the local in the calculation process [33-37].

4 Results and Discussion

4.1 Experimental Design and Environment

Because of the large demand for shared bicycles, the amount of data obtained is also large. Therefore, it is first placed in Hadoop's distributed file system HDFS, and runs under Intel (R) Core (TM) i7-3520M CPU @2.90GHz 2.90 GHz processor, 64-bit operating system. The number of iterations n is 500 and the initial centroid m is 10. Initialized cluster centers are obtained by Canopy algorithm, and the distance between them and K is calculated. Then K-means algorithm can be used to obtain the final clustering results.

4.2 Dataset Collections and Data Pre-processing

Since 2016, shared bicycle has developed rapidly in China. Shared bicycle is widely used as the most ideal means of transportation between many vehicles and people' s residence. The data used are shown in Fig. 6 and Fig. 7.

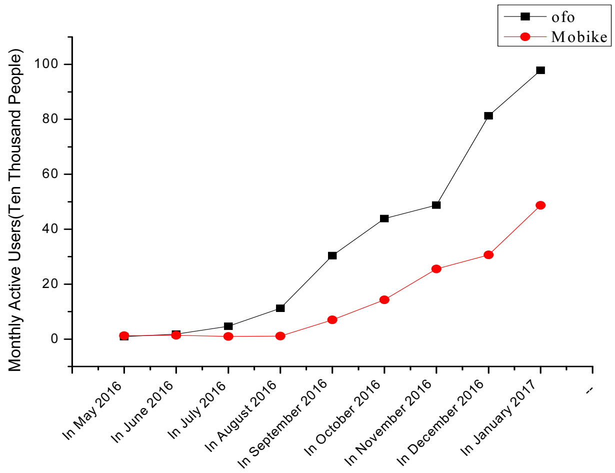

Fig. 6. Comparison of monthly and daily average active users of ofo and Mobike bicycles

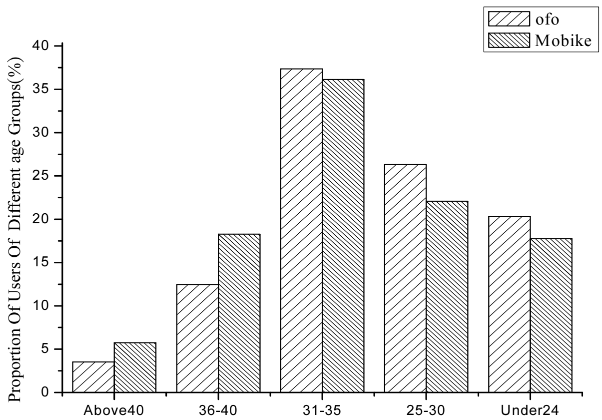

Fig. 7. Percentage of users in different age groups

According to the data obtained, the majority of people using shared bicycles are just commuting within 5 km, as shown in Fig. 8 and Fig. 9.

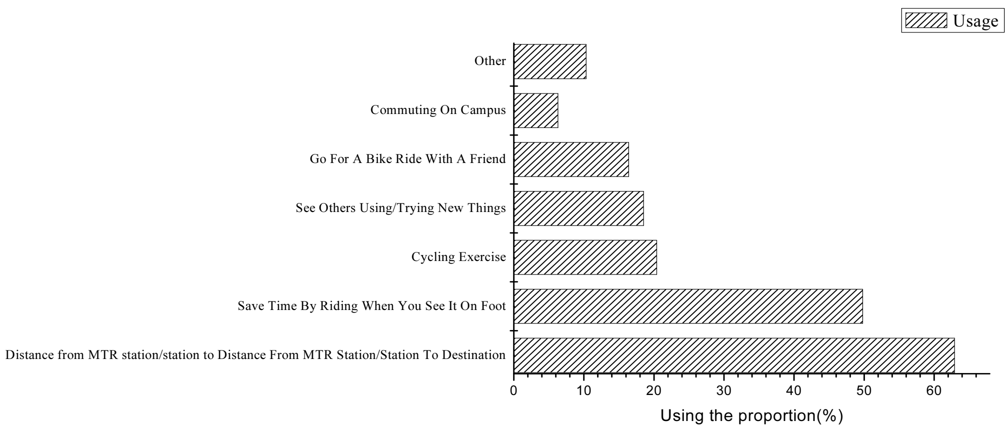

Fig. 8. Shared bicycle usage scenario

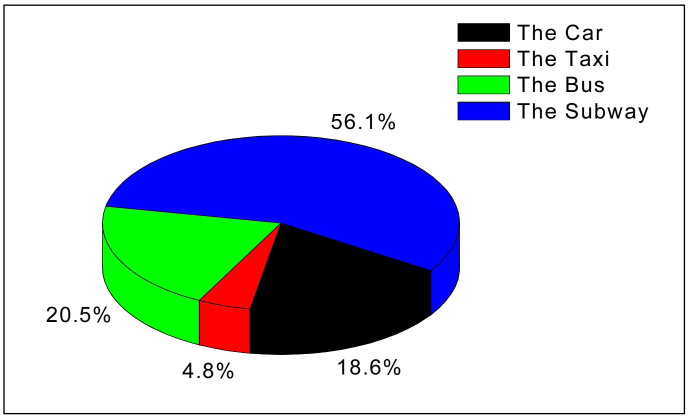

Fig. 9. Shared bicycle travel potential tables for connection or conversion of different traffic modes (origin: AMAP traffic data)

It can be seen from Fig. 8 that the use of shared bicycles is more to facilitate people to complete the “last kilometer” of urban traffic, which also lays the foundation for the layout of shared bicycles. With the increase of the number of users, the relevant data of ofo is chosen to analyze the user’ s usage in a week, as shown in Fig. 10 and Fig. 11.

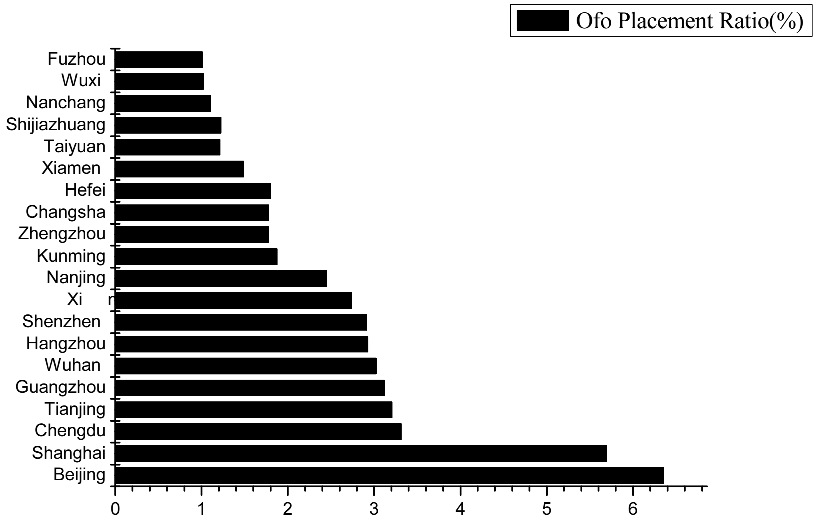

Fig. 10. TOP20 city in the proportion of ofo in 2017 (origin: AMAP traffic data)

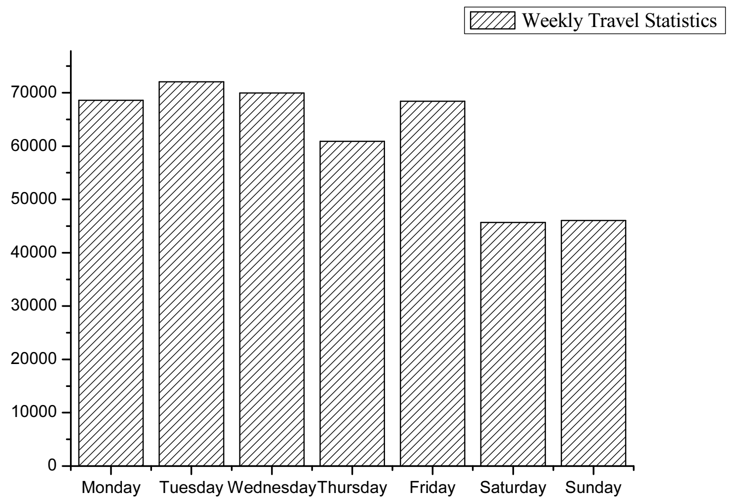

Fig. 11. Weekly travel statistics (origin: AMAP traffic data)

Fig. 10 shows that the number of people go out is the largest on Tuesday, followed by Wednesday, Monday and Friday. This also shows that reasonable prediction of the use demand of shared bicycles during the working day and reasonable scheduling of shared bicycles play a certain role in meeting users' travel needs and alleviating urban traffic.

4.3 Performance Evaluation and Discussion

Canopy clustering is used to optimize the initial clustering center, and the distance between the data points and the cluster center is obtained, from which the shortest distance is selected. In this way, the computational efficiency is greatly improved, and the data is relatively independent at this time. It can be analyzed by using Hadoop platform. From the results of many literatures, it can be seen that the combination of Canopy algorithm and K-means algorithm can improve the efficiency of the algorithm. More importantly, Canopy algorithm improves the accuracy of clustering after optimizing the initial points, and has a larger scope of application. Therefore, it is feasible to combine Canopy algorithm with K-means algorithm for clustering analysis.

In this study, the use of shared bicycles in Beijing is analyzed, and the starting point distribution map of shared bicycle travel potential in Beijing' s Fifth Ring Road is obtained, as shown in Fig. 12.

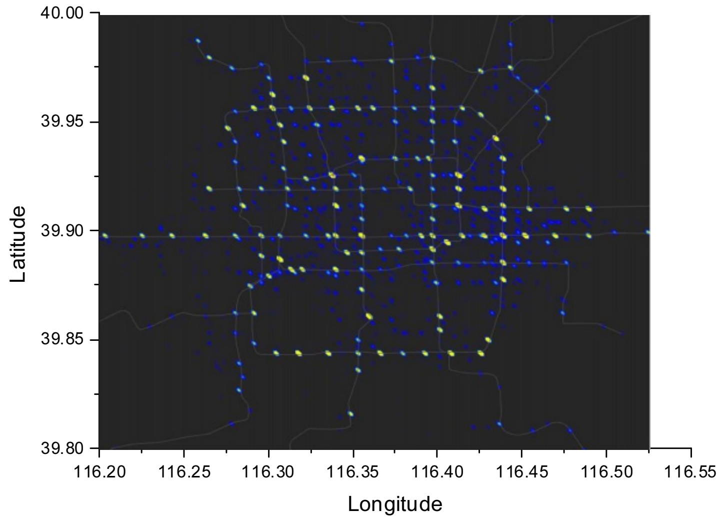

Fig. 12. Distribution map of starting potential of shared bicycle travel in Beijing's Fifth Ring

Note: The grey line is a metro line, and the color of the dots changes from darker to lighter, indicating that the trip volume changes from low to high.

From Fig. 12, it can be seen that the places with great travel potential are mainly along the Metro lines as well as bus station, which also reflects the different capacity of sharing bicycle trips because of different modes of transportation. Among them, the demand for sharing bicycle trips connected or converted to the Metro accounts for 56.1%.

According to the cycling navigation data provided by AMAP, the cycling activities around the metro are concentrated during the peak period in the core area of Beijing. The demand for sharing bicycles is larger in some large stations and areas with more concentrated office buildings, and the demand is more concentrated in the main urban area. The surrounding metro lines are radial, and there is a greater demand in some places.

Through the above analysis, it can be known that the parking location of shared bicycles needs to be laid out according to people's travel conditions. Therefore, a demand model is established by using the hybrid algorithm of K-means and Canopy.

Firstly, Canopy algorithm is used to initialize the data. The result is used as the number of clusters and initial cluster centers required by K-means algorithm. Then K-means clustering is performed. The specific flow of this model is as follows. Firstly, according to the actual situation of demand points, two thresholds are set, in which T1 is the maximum distance between demand points and T2 is the maximum range of each demand point. Secondly, the number and location of demand points are obtained by Canopy algorithm. Thirdly, this demand point is filtered, and the outliers with less demand are deleted to get new data sets. Fourthly, the number of requirement points satisfying the conditions is taken as K value of K-means algorithm, and the location of each requirement point is taken as cluster center. After several iterations, the final clustering results are obtained. Considering the large amount of shared bicycle data, Hadoop parallelization method is used to solve the hybrid algorithm. The distribution of data points before and after clustering is shown in Fig. 13 and Fig. 14.

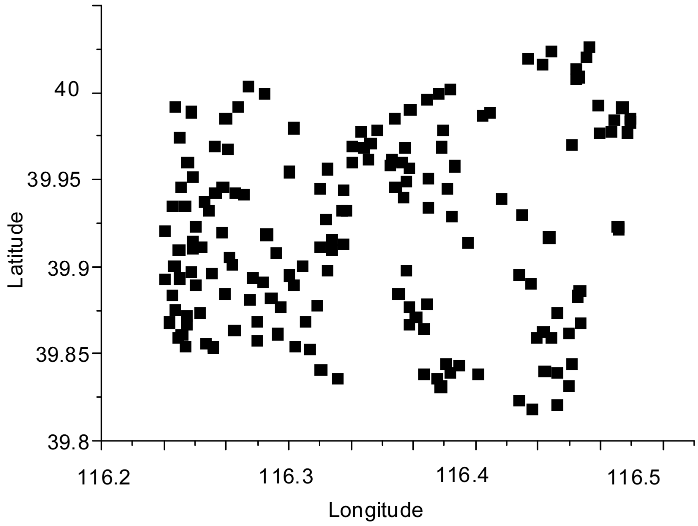

Fig. 13. Data point analysis before clustering

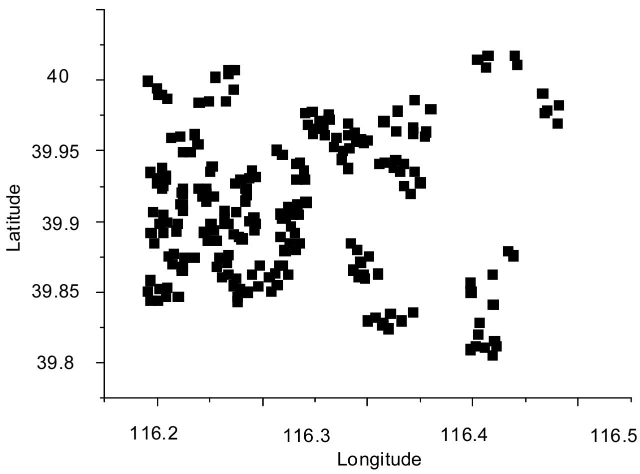

Fig. 14. Data point analysis after clustering

By comparing Fig. 13 and Fig. 14, it can be seen that the clustering of demand points in Fig. 13 is more obvious. Therefore, by mapping the clustering results, it can be found that more shared bicycles need to be arranged in the subway station and bus station to meet the last kilometer of users' traffic trip, while bicycles are arranged in residential areas and companies, which can greatly facilitate users' travel. It also shows that using clustering algorithm to analyze the relevant data is helpful to analyze the layout of shared bicycles.

Beijing has also used "three policies and one phenomenon" in view of the congestion situation in the core urban areas, namely, the new policy of network car appointment, the policy of industry relief, the policy of wall-opening, and the emergence of shared bicycles. In the aspect of shared bicycles, the layout of shared bicycles can be further laid out through the use of people, which can better help people achieve the "last" of public transport. At the same time, congestion in Beijing' s core urban area drops by 6%, effectively alleviating the traffic situation in the area, which also reflects that the demand for sharing bicycles in residential, subway stations, office areas is larger. In the latter layout, these regions need to be given priority consideration.

In summary, using K-means and Canopy algorithm for clustering analysis of demand points can effectively remove some outliers, and can obtain the demand points of shared bicycles, so as to establish a reasonable layout. Doing a good job in the connection or conversion between means of transport can improve the congestion situation of urban traffic to a certain extent. Residence, subway and public areas have become areas with large demand for shared bicycles, which is the key consideration in the implementation of the layout. It can also establish an effective transportation network to make rational use of resources, realize low-carbon travel, and ultimately realize the high efficiency of urban traffic.

5 Conclusion

In this study, the open source data of relevant websites are used to obtain the demand data of shared bicycles in urban traffic network. Through the analysis of Canopy clustering algorithm and K-means algorithm, it is found that there is a large demand in residential areas, subway crossings, bus stop signs and companies. However, the large accumulation of shared bicycles also has a certain impact on urban traffic. Therefore, clustering analysis of these high demand points is needed to remove some outliers, so as to obtain more effective demand and establish a layout model that can meet more users. However, in the process of research, since the weather factors are not taken into account, the layout model may be affected to some extent. Therefore, this factor will also be taken into account in future research, so as to obtain a more comprehensive and reliable layout model. Open source data is used to analyze the demand for shared bicycles in traffic network, and discuss the layout of shared bicycles, which provides some ideas for future operation of shared bicycles, and also provides some theoretical basis for future research on shared economy, so as to realize the low energy consumption and high efficiency development of urban traffic, as well as the rapid development of China's economy.

Acknowledgements

This article was supported by the National Key Research and Development Plan "Mine Safety Situation Analysis, Forecast and Early Warning System" (No. 2017YFC0804408), the National Social Science Major Project subject to tender optimization control of pollutants emission in the era of "big data research" (18ZDA052), the Special Fund of Basic Scientific Research Operating Expenses of Central Universities and Colleges funded the "study on the coupling dual-driven development of export-oriented economy and blue economy in coastal economic areas based on the" One Belt And One Road "strategy" (NO.2017WA03), China Postdoctoral Science Foundation on the project "traffic flow based on cellular automata chaotic scene simulation and the induction strategy research" (No. 202015M581902), and Postdoctoral Scientific Research Project supported by plan of jiangsu province "zero information derived value measurement of the dynamic modeling and simulation research" (NO. 1402067C).

References

[1] H. Jin, F. Jin, J. Wang, Competition and cooperation between shared bicycles and public transit: a case study of Beijing, Sustainability 11(5)(2019) 1-13.

[2] K. Newbold, D. Scott, Insights into public transit use by Millennials: the Canadian experience, Travel Behaviour & Society 11(2018) 62-68.

[3] C. Yang, N. Iqbal, B. Hu, Targeted metabolomics analysis of fatty acids in soybean seeds using GC-MS to reveal the metabolic manipulation of shading in the intercropping system, Analytical Methods 9(14)(2017) 2144-2152.

[4] J.B. Griswold, Y. Mengqiao, F. Victoria, G. Offer, J.L. Walker, A behavioral modeling approach to bicycle level of service, Transportation Research Part A: Policy and Practice 116(2018) 166-177.

[5] Z. Cen, J.D. Schmöcker, A Markovian model of user adaptation with case study of a shared bicycle scheme, Transportmetrica B: Transport Dynamics 7(1)(2017) 223-236. DOI:10.1080/21680566.2017.1378599.

[6] S. Amini, S. Toms, Accessing capital markets: Aristocrats and new share issues in the British bicycle boom of the 1890s, Business History 60(2)(2018) 231-256. DOI:10.1080/00076791.2017.1310196.

[7] A. Fielbaum, S. Jara-Diaz, A. Gschwender, Optimal public transport networks for a general urban structure, Transportation Research Part B: Methodological 94(2016) 298-313

[8] Z. Liu, Application of agent technology in urban traffic signal control system, Digital Communication World 8(2019) 226.

[9] L. Han, Y. Luo, Analysis of shared economy- taking shared bicycle as an example, Modern Marketing (Information Edition) 10(2019) 46-47.

[10] X. Cao, Research on the influence of shared bicycles on urban traffic travel based on the view of slow traffic- taking Lanzhou city as an example, Smart City 5(17)(2019) 22-23.

[11] W.L. Al-Yaseen, Z.A. Othman, M.Z.A. Nazri, Multi-level hybrid support vector machine and extreme learning machine based on modified K-means for intrusion detection system, Expert Systems with Applications 67(2017) 296-303.

[12] H. Liu, J. Wu, T. Liu, Spectral ensemble clustering via weighted K-means: theoretical and practical evidence, IEEE Transactions on Knowledge & Data Engineering 29(5)(2017) 1129-1143.

[13] M. Jaroš, P. Strakoš, T. Karásek, Implementation of K-means segmentation algorithm on Intel Xeon Phi and GPU, Advances in Engineering Software 103(C)(2017) 21-28.

[14] J.F. Tong, User clustering based on Canopy + K-means algorithm in cloud computing, Journal of Interdisciplinary Mathematics 20(6-7)(2017) 1489-1492.

[15] S.C. Popescu, T. Zhou, R. Nelson, Photon counting LiDAR: an adaptive ground and canopy height retrieval algorithm for ICESat-2 data, Remote Sensing of Environment 208(2018) 154-170.

[16] N. Clarke, F. Li, S. Furnell, A novel privacy preserving user identification approach for network traffic, Computers & Security 70(2017) 335-350.

[17] X. Zhang, S. Mahadevan, A bio-inspired approach to traffic network equilibrium assignment problem, IEEE Transactions on Cybernetics 48(4)(2018) 1304-1315.

[18] Y. Yan, S. Zhang, J. Tang, Understanding characteristics in multivariate traffic flow time series from complex network structure, Physica A Statistical Mechanics & Its Applications 477(2017) 149-160.

[19] D. Wang, W.W. Che, H. Yu, Adaptive pinning synchronization of complex networks with negative weights and its application in traffic road network, International Journal of Control Automation & Systems 16(2)(2018) 782-790.

[20] R. Kumar, R. Kumar, K. Lu, Local search methods for k-means with outliers, Proceedings of the VLDB Endowment 10(7)(2017) 757-768.

[21] H. Wang, Q. Wang, W. Wang, Text mining for educational literature on big data with Hadoop, in: Proc. 2018 IEEE International Conference on Smart Cloud (SmartCloud), 2018.

[22] K.Y. Chang, T.P.U. Kyaw, S.H. Chen, The importance of carbon-nitrogen biogeochemistry on water vapor and carbon fluxes as elucidated by a multiple canopy layer higher order closure land surface model, Agricultural & Forest Meteorology 259(2018) 60-74.

[23] W. Wang, Y.Q. Xu, Empirical study on spatial type and functional location of logistics parks in Hebei province based on cluster analysis, Value Engineering 37(34)(2018) 22-23.

[24] S.Y. Gao, N. Li, W. Zhang, Empirical research on regional spatial layout optimization under the perspective of circle economy, Science & Technology Progress and Policy 34(3)(2017) 31-36.

[25] Z. Wang, Y. Sun, Y. Zeng, Substitution effect or complementation effect for bicycle travel choice preference and other transportation availability: Evidence from US large-scale shared bicycle travel behaviour data, Journal of Cleaner Production 194(2018) 406-415.

[26] Y. Li, B. Shuai, Origin and destination forecasting on dockless shared bicycle in a hybrid deep-learning algorithms, Multimedia Tools and Applications (2018) 1-12. DOI:10.1007/s11042-018-6374-x.

[27] X. Liang, G. Si, L. Jiao, Z. Li, Recycling scheduling of urban damaged shared bicycles based on improved genetic algorithm, International Journal of Logistics Research and Applications 22(6)(2019) 519-532. DOI:10.1080/13675567. 2018.1438378.

[28] B. Rakel, Children's influence on dual residence arrangements: exploring decision-making practices, Children and Youth Services Review 91(2018) 105-114.

[29] A. Faghih-Imani, R. Hampshire, L. Marla, N. Elurud, An empirical analysis of bike sharing usage and rebalancing: Evidence from Barcelona and Seville, Transportation Research Part A: Policy and Practice 97(2017) 177-191.

[30] Y. Tomita, A. Nakayama, Demand and cost structure analyses on Japanese successful bicycle sharing system called Ekirin-kun to install cycle ports at railway stations, Transportation Research Procedia 25(2017) 3412-3420.

[31] J. Lin, N. Schofield, A. Emadi, External-rotor $6-10$ switched reluctance motor for an electric bicycle, IEEE Transactions on Transportation Electrification 1(4)(2015) 348-356.

[32] J.B. Griswold, Y. Mengqiao, F. Victoria, G. Offer, J.L. Walker, A behavioral modeling approach to bicycle level of service, Transportation Research Part A: Policy and Practice 116(2018) 166-177.

[33] A. Faghih-Imani, R. Hampshire, L. Marla, N. Eluru, An empirical analysis of bike sharing usage and rebalancing: evidence from barcelona and Seville, Transportation Research Part A: Policy and Practice 97(2017) 177-191.

[34] Strauss, Jillian, L.F. Miranda-Moreno, Speed, travel time and delay for intersections and road segments in the Montreal network using cyclist Smartphone GPS data, Transportation Research Part D: Transport and Environment 57(2017) 155-171.

[35] M. Kaspi, T. Raviv, M. Tzur, Bike sharing systems: user dissatisfaction in the presence of unusable bicycles, IIE Transactions 49(2)(2016) 144-158.

[36] Y. Yan, Y. Tao, J. Xu, S. Ren, H. Lin, Visual analytics of bike-sharing data based on tensor factorization, Journal of Visualization 21(4)(2018) 1-15.

[37] K.B. Campbell, C. Brakewood, Sharing riders: how bikesharing impacts bus ridership in New York city, Transportation Research Part A Policy and Practice 100(2017) 264-282.