Predicting Shared-Bike Routes with Geographic Information System and LSTM Algorithm

Hanfeng Wang[1], Liangbo Zhang[2(✉)], and Ge Zhan[3]

[1] Division of Science and Technology, Beijing Normal University-Hong Kong Baptist University (BNU - HKBU United International College) Zhuhai, Guangdong, China
[2] School of Economics and Management, Harbin Institute of Technology (Shenzhen), Shenzhen, China
`20b957003@stu.hit.edu.cn`
[3] AI Data Analytics Lab, Beijing Normal University-Hong Kong Baptist University (BNU - HKBU United International College) Zhuhai, Guangdong, China

Abstract. Shared bikes in red, orange, yellow and blue can now be seen everywhere in China's first-tier cities, Including Beijing, Shanghai, Guangzhou and Shenzhen. During the rush hour, because a large number of bikes are riding away, there is little reverse inflow, resulting in tidal dilemma. We collected over one million shared bike tracks and conducted data mining with our geographic information system to find out the rules of scientific parking of shared bikes. This project adopts a two-layer LSTM (Long and Short memory neural network) algorithm with spatial-temporal big data from Xiamen.

Keywords: Shared bike · information system · algorithm · neural network

1 Introduction

Buses, subways, taxis, and private cars cannot meet people's urgent needs for short-distance travel, and the "municipal public bicycles" in some cities have not been able to address this issue due to a series of problems such as insufficient bicycle delivery and inconvenience in returning vehicles at particular locations. Based on this, OFO put forward the concept of "sharing economy + intelligent hardware", and created the dockless shared bicycle travel solution, which not only effectively solved the difficulty of urban travel, and saves time and cost for users [12]. Shared bicycles also promote green and low-carbon travel, improve urban congestion, and make cities more beautiful.

The geographical advantage makes the public bicycle system widely used in tourism [13]. Lyon, once the second largest city in France (now ranked third after Marseille), is located in the intersection of central and southern France. It is the only way from northern Europe to Provence and the Mediterranean, a famous scenic spot in southern France. Every year, especially in summer, tourists from northern France, the United Kingdom, the Netherlands, Belgium, and northern Germany flock to Lyon, making it an important tourist city in France. As the second largest exposition center and cultural city in France, Lyon also attracts nearly 6 million tourists every year, second only to Paris,

which is the city with the largest number of foreign tourists per year in France. Lyon's public bicycle station covers almost all the important tourist attractions in the city, and its superior location makes it convenient for users to store and pick up bicycles. That is why Lyon's public bicycles are popular with tourists who want to explore the city.

The temperature in Xiamen is very suitable for outdoor activities such as cycling, and the average minimum temperature throughout the year is above 15 °C, which is highly dependent on temperature. In terms of outdoor sports, except for the relatively bad temperature in December, January and February, the rest of the time is suitable for using bicycles as a means of transportation. The climate of Xiamen does not have obvious characteristics of the rainy season and continuous snowfall in winter, so the suitable climate characteristics make it more convenient for residents to use bicycles as a means of transportation [5].

The overall scale of Xiamen is relatively small in China. The straight-line distance between the two ends of the city is much smaller than that of large and medium-sized cities. This makes the necessary travel radius of residents' daily life smaller, which is precisely the advantage of the urban public bicycle transportation system. It is very convenient for residents to travel directly from the starting point to the destination by bicycle. The terrain of the entire city is very flat, and there are no steep slopes. Its public bicycle allows users to adjust the speed and difficulty of travel according to their own conditions. Generally speaking, a simple shifting device can make it easy for users of different genders and ages to face various urban road conditions. In general, Xiamen's urban scale and terrain characteristics make it more comfortable and convenient for people to use bicycles as a means of transportation.

Shared bikes in red, orange, yellow and blue can now be seen everywhere in China's first-tier cities, Including Beijing, Shanghai, Guangzhou and Shenzhen. In the morning rush hour, a large number of shared bikes gather around the office buildings. In the evening rush hour, a large number of shared bikes are silted up near subways and bus stations. In fact, the tidal phenomenon just deviates from the actual needs of users: in the morning rush hour, users' main demand is to find a car near the subway station [2, 3]. During the evening rush hour, find a bus near the office building and return to the subway station. In the early peak period, shared bikes can also meet the needs of office workers. But in the middle and late peak period, because a large number of bikes are riding away, there is little reverse inflow, resulting in tidal dilemma. We collected data and conducted data mining to find out the rules of scientific parking of shared bikes.

2   Materials and Methods

2.1   Data Collection

First, for this task, we downloaded a dataset from Datacastle, which contains over one million shared bike tracks, as well as the geographical distribution of temples, churches, ancient buildings and bus stops in Xiamen, for subsequent cross-analysis.

Address:
https://challenge.datacastle.cn/v3/cmptlist.html.
Datasets:
shared_bikes_20201221.csv (shared bike tracks)
Religious sites location.csv
Bus stop location.csv

2.2 Data Preprocessing

We use data science and statistical software such as Python and Stata to preprocess the data, and then conduct data mining and visualization. Results are shown in Fig. 1.

We formatted the 5 consecutive days of bicycle trajectory data (positioning data generated by shared bicycles) into gpx format data, and then used the GIS map data browser to display the frequency of bicycle riding in the data coverage area. The interface of this software is simple and powerful, and it is a lightweight ArcMap. Green points represent locations with a frequency greater than 500 and blue ones represent location with a frequency less than 500.

We found that the shared bike track at 8 o'clock near the bus station was the most frequent. We will build a model to verify it later. At the same time, we also conducted correlation analysis and found a greater correlation between ride frequency and longitude of shared bikes (Table 1).

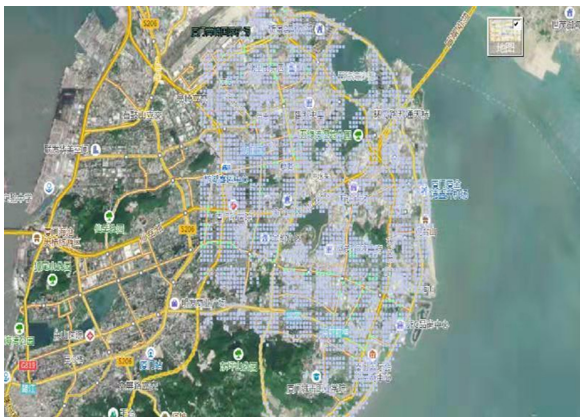The longitude and latitude data of the other two data sets retain three decimal places. In order to merge the subsequent data sets, we take three decimal places for the longitude and latitude data of shared bikes. The data set now has the same number of digits of latitude and longitude. Then we can apply 'merge' to merge two datasets, default to inner join. Finally, we modified the merged data set slightly, excluding some irrelevant data classes, only preserving latitude and longitude and time.

The longitude and latitude data of the other two data sets retain three decimal places [6]. In order to merge the subsequent data sets, we take three decimal places for the

Fig. 1. Riding frequency distribution (green: frequency > 500, blue: frequency < 500).

Table 1. Correlation Analysis.

	freq	hour	longitude	latitude
freq	1			
hour	0.072	1		
longitude	**0.229**	0.109	1	
latitude	0.089	0.063	0.690	1

longitude and latitude data of shared bikes. Finally, we modified the merged data set slightly, excluding some irrelevant data classes, only preserving latitude and longitude and time.

2.3 Algorithm

This project adopts a new RNN (Recurrent Neural Networks) algorithm, two-layer LSTM (Long and Short memory neural network). LSTM is a type of neural network that contains LSTM blocks. LSTM blocks are also called intelligent network units because they can remember values of varying lengths of time. There is a gate in the block that determines whether 'input' needs to be remembered and output 'output'.

Since the riding of shared bikes is a continuous process, the data set can be understood as a 6–9 o'clock time series. LSTM is especially suitable for the prediction and classification tasks of time series type. Therefore, different from the traditional algorithm model [e.g., 7] our strategy is to regard cycling as an event within a period of time and use LSTM algorithm for time series analysis. Time series predictive analysis is to predict cycling characteristics in the future by using cycling characteristics in the past period of time. Different from the prediction of regression analysis model, the time series model is dependent on the sequence of events, and the input results of the model are different after changing the sequence of values with the same size.

3   Results and Discussion

3.1   Evaluation

Integration of shared bike tracks and bus station data

$$MSE = 0.41, RMSE = 0.63 \tag{1}$$

Integration of shared bike tracks and religious place data

$$MSE = 0.91, RMSE = 0.95 \tag{2}$$

Both models achieve ideal RMSE values ≤ 1, and the accuracy of model 1 (Integration of shared bike tracks and bus station data) is slightly better than that of model 2 (Integration of shared bike tracks and religious place data).

3.2   Visualize the Prediction

As an emerging model of the sharing economy, shared bicycles have experienced explosive growth in a very short period of time. Due to the lag of the law, it is difficult for the law to regulate the traditional economic model to solve the problems caused by the current sharing model [9, 11]. For example, the problem of indiscriminate parking of shared bicycles needs to be regulated urgently [10]. Many users park their bicycles at random on streets and bustling areas in the city after riding, which not only hinders normal traffic, but also causes chaos in the appearance of the city and brings hidden dangers to public security [8]. In addition, regulations on shared bicycle riding regulations need to be introduced urgently. Due to the separation of the right to use and ownership of shared bicycles, it is urgent to answer questions such as how to ensure the safety of the bicycle, limit the age of the riders, user insurance coverage, and who is responsible for the accident and claims [1]. The French company JCDecaux, operator of the Lyon bikeshare system, initially put in more than 3,000 specially made bikes and more than 350 bike access stations. Among them, the vast majority of bicycle access points are located in important urban transportation hubs, such as subway stations, railway stations, major bus transfer stations, commercial centers, schools, parks and large and medium-sized residential areas. On the premise of not hindering the existing city facilities, the Xiamen government should strive to meet the requirements of its operators for the planning of access stations for the city's public bicycle system, so that users can more easily pick up or store public bicycles. This would greatly accelerate the awareness and acceptance of urban public bicycle systems by urban residents, and also enables operators to achieve their expected goals as soon as possible, and lays the groundwork for subsequent maintenance and increased investment [4, 14] (Figs. 2 and 3).

The major cities in China are relatively large and cannot be fully organized when setting up public bicycle stations. The entire urban area should be covered with a network structure and densely distributed in major transportation hubs. For the distribution of shared bicycle stations in large and medium-sized cities in China, it is recommended to adopt a radial structure for densely populated areas near major transportation hubs (such as bus transfer stations and subway stations, schools, commercial centers, etc.).
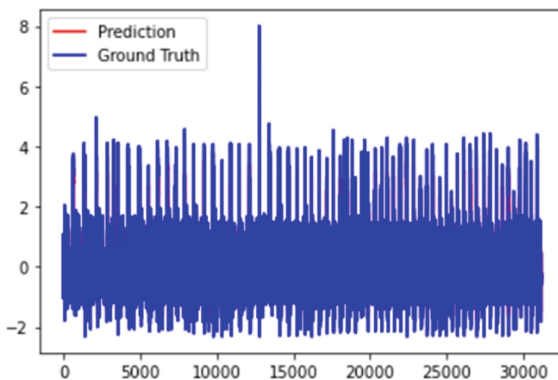

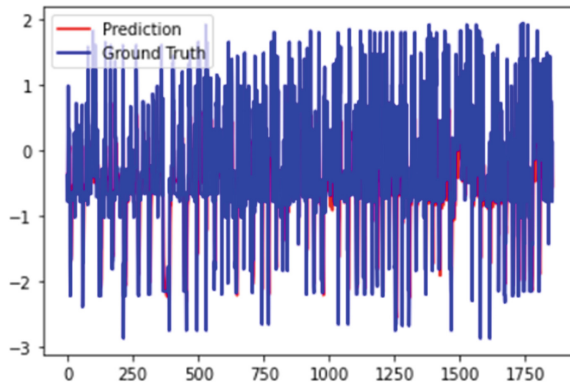**Fig. 2.** Model 1 (Integration of shared bike tracks and bus station).

Fig. 3. Model 2 (Integration of shared bike tracks and religious place data).

Xiamen's bus and subway systems are more similar to other cities. The relationship between the public bicycle system in large cities and the urban bus and subway systems should be more inclined to mutual aid and complementary partners. Xiamen's climate makes public bicycles relatively easy for residents to accept, while cities in relatively cold regions should consider whether the bicycles are stored indoors or with special wind and rain shielding devices, and how to enable users to choose public bicycles for convenient travel in relatively bad weather. When setting up public bicycle systems in different cities, it is necessary to pay attention to these objective factors and take corresponding effective measures to overcome them.

Excellent management mode and high-tech application can make the public bicycle system not only prevent malicious behaviors such as theft and damage, but also speed up the rapid adjustment of the number of shared bicycles, thereby ensuring the fast and good operation of the entire system. Due to the high density of residents in Xiamen, the number of shared bicycles in the station during certain peak hours cannot meet the needs of users. For example, some customers install it on their vehicles without permission to ensure that they can pick up a bicycle the next day. In order to reduce the occurrence of such bad behaviors, the shared bicycle system can launch a service project of online reservation in advance, and allow other users to rent after the reservation time is exceeded, so that operators can plan and schedule the number of shared bicycles at each station as soon as possible.

4   Conclusion

Bus stops and religious places are both important factors influencing shared bikes routes. At 8 o'clock in the morning, there is a serious morning rush near the bus stops, but it's not obvious near the religious places. Near the bus station, due to the obvious morning rush hour phenomenon, parking spots and shared bike allocation can be increased. Citizens are encouraged to park near the bus stop and will be given a discount for riding fees. This project uses innovative algorithms to verify the role of eigenvalues of bus stops and religious places in predicting cycling. The RMSE value of model accuracy is ideal.

The urban public bicycle system in Xiamen, China, is one of the most successful cases in the green and low-carbon transportation system. With the promotion of the concept of "low carbon life" in China, other cities have also introduced shared public bicycle systems, but they have encountered many difficulties and obstacles. Xiamen's public bicycle system has achieved great success in establishing a green public transportation system based on the characteristics of the city, residents' awareness, and external objective conditions. Studying the success factors of Xiamen's shared bicycle system is of great significance for promoting the establishment and development of green transportation system.

References

1. Angeloudis P, Hu J, Bell MG (2014) A strategic repositioning algorithm for bicycle-sharing schemes. Transportmetrica 10(8):759–774
2. Chu J, Duan Y, Yang X, Wang L (2021) The last mile matters: impact of dockless bike sharing on subway housing price premium. Manage Sci 67(1):297–316
3. Gimon D (2018) Bike commuters contribution to balance shared bike systems during peak load. In: 2018 IEEE ISC2, pp 1–2
4. Guo Y, Zhou J, Wu Y, Li Z (2017) Identifying the factors affecting bike-sharing usage and degree of satisfaction in Ningbo, China. PLoS ONE 12(9):e0185100
5. Hu Z, Liu X (2019) The impact of bike sharing on urban transportation. Front Soc Sci Technol 1(4)
6. Jia W, Tan Y, Li J (2018) Hierarchical prediction based on two-level affinity propagation clustering for bike-sharing system. IEEE Access 6:45875–45885
7. Liu Y, Jia R, Xie X, Liu Z (2018) A two-stage destination prediction framework of shared bicycles based on geographical position recommendation. IEEE Intell Transp Syst Mag 11(1):42–47
8. Sun L, Zhou X, Sun Z (2019) Improving cycling behaviors of dockless bike-sharing users based on an extended theory of planned behavior and credit-based supervision policies in China. Front Psychol :2189
9. Tang G, Keshav S, Golab L, Wu K (2018) Bikeshare pool sizing for bike-and-ride multimodal transit. IEEE Trans Intell Transp Syst Mag 19(7):2279–2289
10. Tormo MT et al (2022) Development of a roadmap for the implementation of a sustainable mobility action plan in university campuses of emerging countries. Front Sustain Cities 3:668185
11. Westland JC, Mou J, Yin D (2019) Demand cycles and market segmentation in bicycle sharing. Inf Process Manag 56(4):1592–1604
12. Xin Z, Liang M, Zhanyou W, Hua X (2019) Psychosocial factors influencing shared bicycle travel choices among Chinese: an application of theory planned behavior. PLoS ONE 14(1):e0210964

13. Yang Y, Jiang L, Zhang Z (2021) Tourists on shared bikes: can bike-sharing boost attraction demand? Tour Manage 86:104328
14. Zhanyou W, Dongmei H, Yaopei Z (2020) How to improve users' intentions to continued usage of shared bicycles: a mixed method approach. PLoS ONE 15(2):e0229458