ISSN: 2688-8645

An Effective Method to Analyze Residents' Travel Behavior based on ArcGIS Software

Zhengji An[1]

[1]Shanghai Maritime University, Shanghai, China.

Abstract

Due to China's large population and limited land resources, it is impossible to use a large amount of land to build roads in urban construction. However, a large number of people from towns and villages flood into cities every year. Therefore, it is particularly important to develop urban public transport.As a new urban public transportation tool, ride-hailing not only helps to improve employment and income levels, but also improves people's travel efficiency and quality. With the characteristics of round-the-clock operation and stop-and-go, ride-hailing is favored by more and more residents.In addition, with the rapid development of geographic information system (GIS) technology and the extensive deployment of global positioning system (GPS), a large number of individuals' mobile trajectory data are widely stored, providing a new idea for the analysis of residents' travel behavior.Based on the application of GPS technology in online ride-hailing vehicles, this paper studies residents' travel behavior, processes OD point data of a certain workday online ride-hailing vehicles in Chengdu, discusses residents' travel behavior from the spatial dimension, and combines it with the actual situation to finally reach a conclusion.

Keywords

OD point data of online ride-hailing, Resident travel, Nuclear density analysis, Point of interest.

1. Introduction

1.1. Research Background

Since China's reform and opening up, the rapid economic development has accelerated the urbanization process significantly. According to the 2019 population census released by the National Bureau of Statistics, the permanent urban population in China will reach 848 million in 2019, and the urban population will account for 60.60% of the total population (the urbanization rate) [1]. Other statistics show that 13 cities in China have a population of more than 10 million, among which Chongqing, Shanghai and Beijing have a total population of more than 20 million. By more than 60% of urbanization rate and urban population in swelling, every year in the city's economic development at the same time, also brought a series of social contradictions that nots allow to ignore, such as public service resources such as education and health care increasingly nervous, city road show "pendulum" traffic flow characteristics, exacerbating the problem of traffic congestion, urban environmental pollution, etc. These problems seriously affect the current management and future planning of cities and bring great challenges to the strategy of sustainable urban development. Among them, urban traffic congestion is not only manifested in residents' morning and evening rush hours, nor is it limited to traffic trunk roads and sub-trunk roads. The period of congestion is constantly expanding, and the congested roads include life trunk roads and even expressways. Urban congestion tends to be normalized. Traffic congestion not only reduces the work efficiency of urban residents,

ISSN: 2688-8645

increases their travel time cost and travel time, but also gradually becomes the main pollution source of the city, affecting residents' physical and mental health.

The main reasons for these contradictions are the rapid growth of the population and a large number of rural population into the city as well as the lag of urban public facilities construction. In order to alleviate traffic pressure, local governments have been committed to improving road conditions by means of transportation projects, and at the same time, they have strengthened the construction of public transportation facilities and reduced the total road area occupied by means of transportation. If residents want to get to the destination of their city quickly, online ride-hailing is undoubtedly an important means of transportation for urban residents, because online ride-hailing has the characteristics of round-the-clock operation and stop-and-go. Therefore, the GPS data of online ride-hailing vehicles contains the travel demand information of urban residents and also reflects the traffic conditions of cities. Moreover, these trajectory data are real-time and easily accessible, making it feasible to mine residents' travel behavior rules, discover residents' travel behavior characteristics and study residents' travel interest points through ride-hailing trajectory data .

1.2. Research Status at Home and Abroad

1.2.1 Research status of floating vehicle trajectory data

In this respect, the research of Europe, Japan, the United States and other developed regions and countries started earlier, they are in the world's leading position. As early as 1986, Europe began to carry out research on intelligent transportation, leading the world in the development and implementation of GPS trajectory data collected by floating vehicles. In the 1970s, Japan began to develop an intelligent transportation information platform (JSK) to provide a basic platform for the subsequent research and development of intelligent transportation information [2]. In 1990, The first road vehicle information and communication system was established in Japan. In the late 1960s, the United States has already begun to study intelligent transportation systems. In the 21st century, the INTELLIGENT transportation information system platform in the United States has been well developed, among which the representative intelligent transportation information systems are as follows: The American urban Transportation Guidance System (TRAVTEK);A floating car system developed by the State of Minnesota in collaboration with ford Motor Research; Floating car experiment established in Florida [3].

1.2.2 Research status of urban residents' travel behavior

Residents travel behavior study began in 1994, the United States was aimed at transportation in medium-sized cities such as Memphis, the trip survey of main access by sampling the family members of one day trip intensity, mode, travel purpose and travel and end points, such as information, sample and through the expansion of the city's estimated total travel. Murakam is the first to introduce GPS technology into the field of travel analysis. He asked 100 volunteers to install a Personal Digital Assistant and a GPS receiver connected to it on their private cars, and obtained the respondents' travel purpose, location and time information within six days. Jean Wolf further discusses how to organize and deal with the GPS data, obtain more travel information, she will travel GPS data of 13 respondents use parking time threshold value division, the use of land use types and travel purpose to infer the corresponding relationship of the participants' travel purpose, confirmed the trip purpose using GPS data acquisition and the feasibility of behavior [4].Deng Zhongwei et al. analyzed the spatio-temporal dynamic evolution law of Shanghai residents' travel activities by processing the one-day track data of 2,2,000 taxis in Shanghai [5].

ISSN: 2688-8645

2. Track Data Processing of Online Ride-Hailing Vehicles

2.1. Data Sources

The data source of this paper is the online ride-hailing order data of Chengdu on November 3, 2016 provided by Gaia Data Opening Plan, which contains nearly 190,000 OD point order data. The original data is stored in the form of Excel, mainly including order number, latitude and longitude, time stamp, starting and ending billing time, travel time and distance as the crow fly. An example of the raw data is shown in Figure 1.

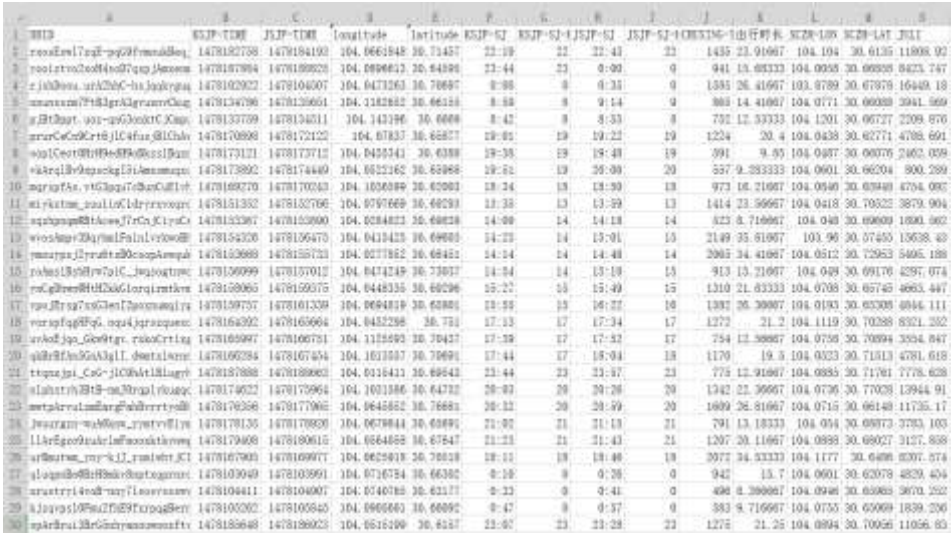

Figure 1. Sample of the raw data section

Where, the specific meaning of each column value is as follows: Column A: order ID, as the unique identification of this order; Column B: Start billing time stamp; Column C: : End billing time stamp; Column D: Longitude coordinates at the beginning of billing; Column E: Latitude coordinates at the beginning of billing; Column F: Start billable time, accurate to minutes; Column G: Start billable time, accurate to hour; Column H: end billable time, accurate to minutes; Column I: End billable time, accurate to hour; Column J: Travel time in seconds; Column K: Travel time in minutes; Column L: longitude coordinates at the end of billing; Column M: Latitude coordinates at the end of billing; Column N: the straight-line distance from O to D.

2.2. Data Preprocessing

The pre-processing of original data in this paper mainly includes the preprocessing of trip time and trip linear distance.During the operation of the online ride-hailing vehicles, data loss or information transmission delay may occur due to the occlusion of high-rise buildings, driving into underground tunnels and other reasons, or the malfunction of the GPS device of the online ride-hailing vehicles.Therefore, data preprocessing is needed before data analysis and map matching, so as to reduce errors.

(1) Travel time screening

In the Excel table, "filter" was performed on column K of travel time. In this paper, data with travel time less than two minutes were screened and deleted.

(2) Selection of straight-line travel distance

Column N, the travel linear distance, is calculated from the latitude and longitude coordinates of OD two points. Therefore, the data with unchanged latitude and longitude coordinates should be deleted. Therefore, in the Excel table, the "filter" operation is carried out on column N to filter it out and delete it.

ISSN: 2688-8645

After these two steps of data preprocessing, the final remaining 188,414 pieces of data, with no missing data such as latitude and longitude coordinates, time stamp and order number of each piece of data, are more accurate and can be used for residents' travel behavior analysis.

2.3. Longitude and latitude data processing based on ArcGIS

In this chapter, the operation of online ride-hailing vehicles in Chengdu during the three periods of morning peak period (7:00-9:00), flat hump period (10:00-11:00) and evening peak period (17:00-19:00) is analyzed. The software-ArcGIS is used to visualize the longitude and latitude coordinates of starting and ending points and conduct kernel density clustering analysis, so as to obtain the spatial characteristics of residents' travel.

Firstly, the coordinates of the starting and ending points of the three time periods were screened out and saved in the format of ".xls". They were named as "morning peak", "flat hump peak" and "evening peak" respectively. The boarding and boarding times of the three time periods were counted respectively, as shown in Table 1. Secondly, open ArcGIS, click "Add Data", import the filtered data into the layer, right-click and click "Display XY data", X field corresponds to longitude and Y field corresponds to latitude, and the starting and end points can be visualized. Click "Add Data" to import the base map of Chengdu Road Network into the layer. Again, export the data of the starting point and the ending point respectively, and save the type as "Shapefile". Finally, select "Density Analysis" under "Spatial Analyst Tools. TBx" in "directory", "system toolbox", double-click "Core density Analysis", and add the file to "input point or broken line elements" as the starting and ending Shape file just saved to get the final clustering result.

Table 1. Statistics of loading and unloading times after pretreatment

	get on	get off	total
morning peak period	17001	14396	31396
flat hump period	10742	11048	21790
evening peak period	23576	23347	46923

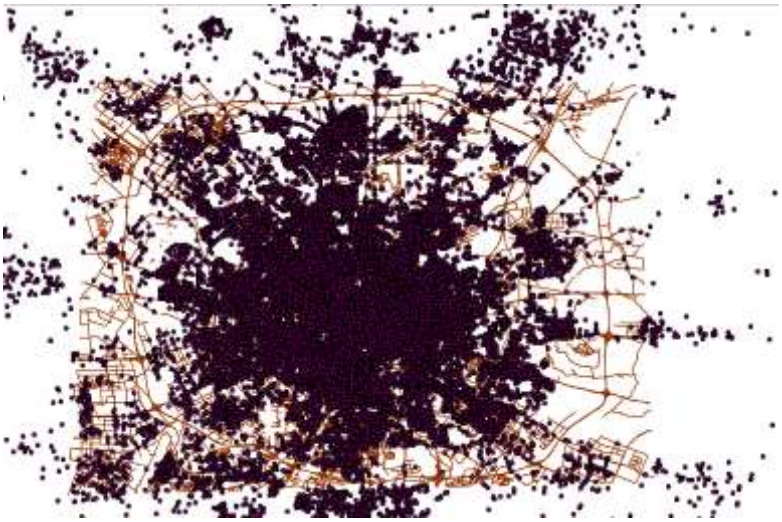

Figure 2. Comparison diagram of Chengdu residents' boarding point and actual road network

The preprocessed one-day data of Chengdu on November 3, 2016 was directly imported into ArcGIS for "display XY data" operation, and the starting and ending points were visualized and imported into the road network base map of Chengdu, as shown in Figure 2, which is the matching map between the starting point, longitude and latitude of Chengdu residents' daily

ISSN: 2688-8645

ride-hailing journey and the road network base map. The base map of Chengdu city road network shows the surrounding and inner areas of Chengdu Circum-city expressway. As can be seen from the figure, the travel location of residents in Chengdu matches with the road network of Chengdu, which meets the research requirements.

3. Residents' Travel Space Analysis based on THE OD Data of Online Ride-Hailing Vehicles

3.1. Operation Steps

Based within the second ring road of Chengdu and around area of morning peak period (7:00-9:00), flat hump period (10:00-11:00) and evening peak period (17:00-19:00) residents travel OD points were analyzed, and observe the trip density distribution within the second ring road, by ArcGIS respectively into three periods of latitude and longitude coordinates, to "show the XY data" operation, after all point visualization import network and reproduction of Chengdu will be imported from ArcGIS online comparison analysis of regional land use type of intensive reproduction of Chengdu.

Notes when displaying XY data:

1. The file must be“.xls”, which is the Excel file of 1997-2003 version, which can be implemented by saving as.
2. XY data (in Excel) each column of data must be a numeric decimal system.
3. Selection of coordinate system: geographic coordinate system.
4. Kernel density analysis can only be performed after data is exported.

Figure 3 shows the distribution of OD points during the morning peak period (7:00-9:00). In Chapter 2.3 of this paper, the operation steps of longitude and latitude data processing and kernel density analysis are explained in detail, which will not be repeated here. The kernel density analysis tool is used to calculate the density of elements in their surrounding areas. When setting the kernel density analysis (Figure 4), note that the "Population field" selects NONE so that each element is counted only once. The unit size of the output pixel size is a smaller value in the width or height of the output range in the output space reference divided by 250.

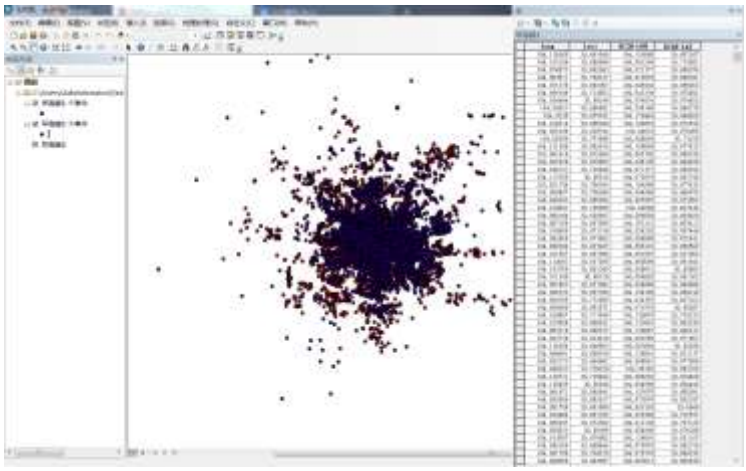

Figure 3. The distribution of OD points during the morning rush hour

ISSN: 2688-8645

Figure 4. Nuclear density analysis Settings

After the base map of Chengdu is imported from ArcGIS Online, it can be found that the kernel density analysis result layer covers the base map of Chengdu road network. Dense areas can only be seen, but the location of dense areas cannot be determined, as shown in Figure 5. Right-click on the kernel density layer in the Content List and select Layer Properties. In the Display bar, select Transparency and adjust the transparency as required, as shown in Figure 6.

After setting the transparency to 50%, not only the core density analysis layer but also the road network base map can be seen, as shown in Figure 7. The color from pink to purple represents the density of OD points, and the darker the color is, the more residents get on/off the bus.

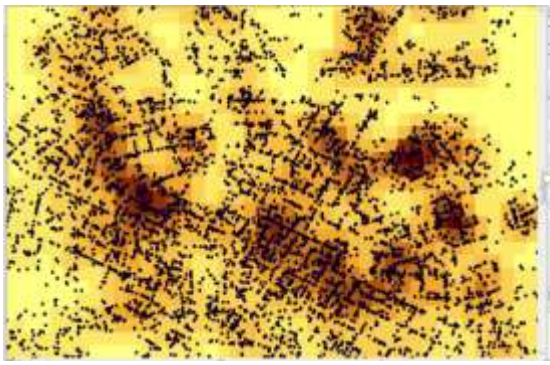

Figure 5. Unable to display base image

Figure 6. "Transparency" adjustment

Figure 7. Analysis of nuclear density in Chengdu under 50% transparency

ISSN: 2688-8645

Through the above operation steps, O and D points of residents' travel during morning peak period (7:00-9:00), flat hump period (10:00-11:00) and evening peak period (17:00-19:00) are respectively analyzed. After obtaining the layer shown in Figure 9, the dark-colored area is compared with the base map of Chengdu city to find out the points of interest and the land type of the points of interest, and analyze the purpose and rationality of residents' travel behavior.

3.2. Nuclear Density Analysis Starting and Ending Points

During the morning peak period, residents take ride-hailing cars mostly in residential areas, some near primary and middle schools, and in areas including companies and medical places. It can be inferred that after taking online ride-hailing cars to the company or sending children to school first, residents take online ride-hailing cars near the school to the company. Of course, some residents need medical treatment or go to the hospital for work, and some take long distance trips to the railway station.

During the morning peak period, the hot spots of the starting and finishing points are all scattered, while during the flat hump period, the hot spots are all concentrated in the downtown area within the first ring road. Residents take online ride-hailing cars from their residential areas to leisure and entertainment venues or to companies; Some residents return to residential areas from hospitals by online car-hailing after completing medical treatment. There is also a lot of government activity in government offices. It can be seen that residents whose starting and ending points are located in entertainment places do not need to go to work, such as women and retirees. Residents whose starting point is located in hospitals may be those residents who arrive at hospitals for medical treatment during the morning peak period, complete medical treatment during the flat hump period and take online taxi to go home.

During the evening peak period, residents frequently take online ride-hailing cars for different purposes. Some residents go home after work, and some go to entertainment places after work. There are residents who travel from residential areas to places of entertainment, and there are residents who travel home from places of entertainment. During the evening peak period, residents may take online ride-hailing cars for multiple trips, such as going from the residential area to the shopping mall and then going home from the shopping mall, or taking a taxi to the entertainment places (such as hotels) after work from the company.

In general, the starting and ending points of residents who travel by online ride-hailing are mostly used by companies, residential areas, entertainment, medical care, government and other land types. This part of residents are mainly adults, which also indicates that taking online ride-hailing requires a certain economic foundation. Residents who travel by online ride-hailing are mainly distributed near and within the Second Ring Road. Residents in this area make frequent and large trips, which also provide drivers with more opportunities to pick up orders and quickly arrive at the location of passengers, thus playing a good guiding role for both passengers and drivers.

4. Conclusion

This paper based on the net about car OD points data analysis for the residents' travel behavior research background, on the basis of studying process, and make full use of network about car residents travel to improve efficiency and quality, has the characteristics of all-weather operations and the stop or go, reference to the current situation in this paper, a way of combining ArcGIS analysis residents travel behavior research, have certain objective authenticity, can give the city planning, infrastructure, etc. To provide the reference. Since there are many factors affecting residents' travel behavior, such as travelers' income, travel distance, age, occupation, gender, etc., these information cannot be obtained from the OD point data of

online ride-hailing vehicles. In addition, the research basis of this method requires GPS data, so it is not suitable for the study of walking, bus, self-driving and other travel modes.

Since the starting and ending points and driving trajectory of online ride-hailing vehicles are completely determined by passengers, the OD point data of online ride-hailing vehicles to a certain extent reflects the interest points of urban residents, the hot sections of online ride-hailing vehicles and the congestion of the sections. Therefore, the guiding significance of this paper includes:

(1) As far as drivers and passengers are concerned

Net about presents a regularity in the process of passenger car, through the summary of the data can be found that the law, to help inexperienced about car drivers looking for potential customers in time to reduce the idle time, increase the operating income, for residents need on the net about car travel also can refer to these POI, in need on the net about car more about car timely find the empty net.

(2) For the administrative departments of urban transportation and land planning

Areas with dense starting and ending points for online ride-hailing reflect, to some extent, places with frequent activities (such as residential areas and railway stations) and high popularity (such as central clusters) in cities. These dense areas obtained through cluster analysis can provide guidance for the construction of public parking lots, commercial residential areas and urban road planning. It is helpful to prevent, find or solve the traffic problems and allocate the urban traffic resources reasonably by excavating the hot spots of residents' travel.

Acknowledgements

Natural Science Foundation.

References

[1] National Bureau of Statistics. China Statistical Yearbook [J]. China Statistical Press, 2010-2019.

[2] Richard Bishop.Floating Car Data Projects Worldwide:A Selective Review.ITS America Annual Meeting, 2004.

[3] Boyce D E, Kirson A, Schofer J L.Design and implementation of ADVANCE:The Illinois dynamic navigation and route guidance demonstration program [C]//Vehicle Navigation and Information Systems Conference.IEEE, 1991:415-426.

[4] Wolf J. Using GPS data loggers to replace travel diaries in the collection of travel data[D]. Georgia: Georgia Institute of Technology, 2000.

[5] Deng Z, J M, et al.Deriving rules for trip purpose identification from GPS travel survey and GTS data:a machine learning approach[C]. Proceedings of the 7th International Conference on Traffic and Transportation Studies,Kunming,China,2010:768-777.