Analysis of Bus Accessibility Based on GIS and Big Data of Crowd -- Taking Wansheng Economic Development Zone as an example

Hai Xiao1*, Xu Liang2, Yuan Zheng2, Xinyi Xia2

1 College of Architecture and Urban-Rural Planning, Sichuan Agricultural University, Dujiangyan, China
2 College of Architecture and Urban-Rural Planning, Sichuan Agricultural University, Dujiangyan, China
* Corresponding author’s e-mail: 41228@sicau.edu.cn

Abstract. Public transportation issues are increasingly becoming a key constraint to the sustainable development of urban economic and social development. Based on crowd big data and GIS spatial analysis methods, this article analyzes the accessibility of the public transportation in Wansheng Economic and Technological Development Zone (hereinafter referred to as the development zone). A comprehensive evaluation and analysis was conducted to study the matching relationship between population density and existing public transportation stations, and to propose the optimal suggestions. The results show that the overall public transportation accessibility in the development zone is average level, and the public transportation accessibility is better in the densest crowded areas, but in the east and southeast of the development zone, crowd activities are poorly matched with the public transportation network, and residents’ travel demands can’t get public transportation with timely responses.

1. Introduction
Urban public transportation greatly meets the needs of urban residents for long-distance travel. However, current public transportation routes are unreasonable, public transportation lines are not standardized, and public transportation efficiency is low. Many scholars have conducted in-depth research on the optimization of public transportation systems. In recent years, there have been more and more researches on public transportation using GIS and big data. Zhang Xuemei et al. researched and analyzed 5 common accessibility measurement models, and analyzed a combined accessibility measurement model for public transportation [1]. Based on Sina Weibo and Baidu map data, Lei Chengcheng studied the dynamic characteristics of crowd activities at different time scales, and obtained the relationships between urban population dynamics and land use types [2]. Liu Song et al. evaluated the accessibility of urban and rural public transport in Haining city based on GIS research, and constructed a Haining urban and rural public transport accessibility evaluation model using the modified spatial barrier model and cumulative opportunity model [3]. At present, more researches focus on large cities and large regions, and less research on small cities and small regions. In order to improve the efficiency of public transportation in small cities further, this article tries to sort out various traffic data in the development zone and combine the crowds. Big data analyze the characteristics of residents’ travel, and evaluates and optimizes the accessibility of public transportation in the development zone.
2. Research objects, data sources and research methods

2.1. Research objects
The development zone now has 42 bus lines, covering 2 streets and 8 towns in the whole district. The cumulative operating mileage is about 350 Km. Some bus line information in the urban area is shown in Table 1.

Table 1. Urban bus lines information.

| Bus routes     | Terminal station          | Operation time | Long(km) | Remarks            |
|----------------|---------------------------|----------------|----------|--------------------|
| 101 routes     | No.49 school A - No.49 school B | 06:30-22:20   | 6.7      | 7-15min/shift     |
| 102 routes     | No.49 school B - No.49 school A | 06:30-22:20   | 6.7      | 5-10min/shift     |
| 103 routes     | CCB - Hong Yan            | 06:30-21:30   | 8.9      | 15min/shift       |
| 104 routes     | VAO - VAO                 | 06:30-21:00   | 6.9      | 15min/shift       |
| 105 routes     | Cultural Center - YaoZiKou| 06:00-22:00   | 4.8      | 5-10min/shift     |
| 107 routes     | CCB - Jianshe Town        | 07:20-18:10   | 6.2      | 5-10min/shift     |
| 108 routes     | YaoZiKou - Cultural Center| 06:30-21:30   | 4.9      | 10min/shift       |
| 131 routes     | YaoZiKou - No.49 school A| 06:30-22:00   | 6.1      | 15-20min/shift    |
| 201 routes     | Cultural Center - YuTianBao| 06:00-22:00  | 5.5      | 7-10min/shift     |
| 202 routes     | Cultural Center - NongLin | 06:40-16:35   | 9.8      | 5-10min/shift     |
| 203 routes     | Jiangshan City - Nantong General Hospital | 06:30-21:30 | 5.5 | 10min/shift |
| 331 routes     | CCB - Nantong             | 06:00-21:30   | 9.0      | 10min/shift       |
| 332 routes     | CCB - Zhilu               | 06:03-21:25   | 7.8      | 10min/shift       |

2.2. Data sources
The data used in this article mainly include: the remote sensing image data of the development zone in 2019 from Google Earth, the 2008-2019 "Chongqing Statistical Yearbook" data from the website of the Chongqing Municipal Bureau of Statistics, the 2013 "Wansheng Economic Technology" from the website of the Development Zone Data from the Urban and Rural Master Plan of the Development Zone, and big data on the population positioning of the development zone in 2020 from Easygo. Due to different data sources, before performing spatial analysis, various types of data need to be converted and unified to the Xi’an 80 coordinate system. Part of the bus station and residential area data are from field research.

2.3. Research methods
The research methods used in this study include kernel density analysis, overlay analysis, buffer analysis and other analytical methods. Kernel density analysis is to calculate the unit density of point and line element measurement values in the specified area, which can intuitively reflect the distribution of discrete measurement values in continuous areas [4]. Overlay analysis is to carry out the various data layers composed of relevant subject layers. Overlay produces a new data layer, which integrates the attributes of the original elements of each layer. Buffer analysis is to establish polygon solid area data within a certain width around it, and this area data can be described vividly the closeness of the point, line and area data to be analyzed in geographic space.

2.4. Determine the evaluation factor
According to the bus line standard in the "Urban Road Traffic Planning and Design Code", combined with the actual situation of the development zone's bus system, four indicators of line network density, average stop distance, non-linear coefficient and repeat coefficient are selected for evaluation [5]. The specific evaluation methods are shown in Table 2:
Table 2. Evaluation factors.

| Evaluation factors                                    | Formula          | Explain                                                                 | Range     |
|-------------------------------------------------------|------------------|-------------------------------------------------------------------------|-----------|
| Density of public transportation network              | \[ \rho_r = \frac{L}{S} \] | S is the area of the constituency, L is the length of the center line of the public transportation line, \( \rho_r \) is the density of the public transportation line network | 3~4 km/km² |
| Average station distance                               | \[ l = \frac{s}{n-1} \] | s is the space distance between the first and last station, n is the number of stations, l is the average station distance | 0.5~0.8km |
| Non-linear coefficient (correcting)                   | \[ \bar{x} = \frac{s}{s_{min}} \] | \( s_{min} \) is the shortest road distance of the first and last station, \( \bar{x} \) is the non-linear coefficient after correction | 1.15~1.2  |
| Repetition factor                                      | \[ R = \frac{L}{L_s} \] | \( L_s \) is the total length of the centerline of the road network, R is the repetition coefficient | \( \leq 5 \) |

The final result is multiplied by the corresponding coefficient and added together to get the final accessibility result. The correlation coefficient is determined comprehensively through relevant literature and questionnaire surveys. The reasonable value range of the accessibility index \( Ac \) is calculated to be 1.35-2.95 according to the standard value of each factor. The calculation formula is as below:

\[
Ac = 0.3\rho_r + 0.1l + 0.35\bar{x} + 0.25R
\]

3. Results and analysis

3.1. Evaluation and Analysis of Public Transport Network

The average stop distance is the ratio of the route distance between the first and last stop to the number of stops, and it reflects the density of bus stops. As shown in Table 3, 12 bus lines, including No. 101 and No. 102, do not meet the standards. Among them, 11 bus lines except No. 103 and No. 107 are too close to each other and need to be appropriately optimized and reduced. 0.5 Km below the specification value.

Table 3. Station distance.

| Line | Terminal station                        | Station number | Station distance (km) |
|------|-----------------------------------------|----------------|-----------------------|
| 101  | No.49 school A - No.49 school B         | 15             | 0.44                  |
| 102  | No.49 school B - No.49 school A         | 14             | 0.48                  |
| 103  | CCB - Hong Yan                          | 8.5            | 1.04                  |
| 104  | VAO - VAO                               | 35             | 0.20                  |
| 105  | Cultural Center - YaoZiKou               | 18             | 0.27                  |
| 107  | CCB - Jianshe Town                      | 10.5           | 0.59                  |
| 108  | YaoZiKou - Cultural Center              | 14             | 0.35                  |
When evaluating the non-linear coefficient of the bus line, considering that the development zone is located between two mountains, in order to modify the influence of the terrain on the research results, the corrected non-linear coefficient is calculated by using the shortest spatial distance between the first and last stations. It can be seen from Table 4 that in the uncorrected state, only one bus line's non-linear coefficient meets the specification, and the non-linear coefficients are all reduced after correction. Therefore, the bus lines in the development zone are relatively tortuous.

| Line | Terminal station | Non-linear coefficient | Non-linear coefficient (correcting) |
|------|------------------|------------------------|--------------------------------------|
| 101  | No.49 school A - No.49 school B | 3.0 | 2.04 |
| 102  | No.49 school A - No.49 school B | 3.0 | 2.06 |
| 103  | CCB - Hong Yan | 1.9 | 1.12 |
| 104  | VAO - VAO | 3.4 | 1.81 |
| 105  | Cultural Center - YaoZiKou | 1.5 | 1.19 |
| 107  | CCB - Jianshe Town | 2.3 | 1.07 |
| 108  | YaoZiKou - Cultural Center | 1.5 | 1.24 |
| 109  | CCB - No.49 school A | 1.4 | 1.11 |
| 108  | Cultural Center - YuTianBao | 1.1 | 1.01 |
| 202  | Jiangshan City - Nantong General | 1.2 | 1.12 |
| 203  | Jiangshan City - Nantong General | 1.8 | 1.62 |
| 331  | CCB - Nantong | 2.0 | 1.10 |
| 332  | CCB - Zhilu | 2.3 | 1.27 |
| Average | | 1.87 | 1.33 |

After the summary calculation, the average value of each index is obtained. As shown in Table 5, the accessibility Ac is 3.35 after the weight calculation, which is higher than the normal range of accessibility, which also indicates the accessibility of the bus network in the central urban area of the development zone. Generally, the main reason is that the density of the bus line network is small, the station spacing is too tight, the lines are relatively tortuous, and the line repetition rate is relatively large.

| Evaluation factors | range | value |
|--------------------|-------|-------|
| \( \rho_r \) | 3~4km/km² | 2.53 |
| \( l \) | 0.5~0.8km | 0.41 |
| \( \bar{x} \) | 1.15~1.2 | 1.33 |
3.2. Crowd Big Data and Public Transport Accessibility Analysis

3.2.1. Crowd big data analysis. The population distribution and the setting of the public transportation stations are mutually influencing and interacting. In most cases, the layout of the public transportation stations should follow the migration changes of population density. As the closest analysis data to the real population distribution data, crowd big data of Easygo is widely used in all aspects of urban planning.

Figure 1 shows that the hotspots of crowd activities in the development zone are highly compatible with residential communities, especially in high-rise building communities and the higher the crowd density. The most active area is concentrated in the old city on the east side of the city. The core area includes the five blocks of Transportation Bureau, Guoneng Tianjie, Yunan Pearl, Pedestrian Street, and Juxinyuan. Areas where the population migration is more obvious are Jiangshan city on the north side of Wanxin Road and Tashan on the south side of Tashan Road.

3.2.2. Transit accessibility analysis. Superimposing the population heat map with the various elements of the public transportation system can visually discover the problems in the public transportation system and provide an objective reference for improving the accessibility of public transportation. Figure 2 shows that there is an inconsistency between the bus passenger transport center and the crowd hotspot area, especially in the hotspot area on the southeast side of the city. The sparse distribution of bus stops is inconsistent with the crowded area on the east side of the city, especially in Wansheng Avenue and Dongcheng Avenue. The sparse distribution of bus stations on both sides of the road does not match the high-density crowd activity area. The planned bus station land is inconsistent with the status quo. In the
hot spot on the east side of the city, the Yongli Palace, the Pearl of Southern Chongqing, and the Jiangshan City on the north side of the city, it is far away from the planned bus station land, which obviously does not meet the needs of dense crowds for public transportation.

Figure 2. Analysis of bus station accessibility.

4. Conclusion
Based on crowd big data and GIS related analysis tools, this paper evaluates and analyzes 4 public transportation accessibility indicators, including the average station distance and non-linear coefficient of the development zone. The results show that the overall public transportation accessibility of the development zone is average. Using crowd big data to analyze the matching of bus stations and passenger transportation center, the results show that the matching degree is poor in the east and southeast of the development zone. Field surveys also show that the travel needs of residents in this area can’t be responded by the bus network in a timely manner. In response to the above problems, optimization and reduction should be made in areas with high bus station redundancy, and bus stations should be added in areas with high crowd density and sparse stations.

References
[1] Xuemei Z. (2014) Research of public transport accessibility on GIS – A case study of Luzhou. Southwest Jiaotong University, Chengdu.
[2] Chengcheng L. (2016) Spatiotemporal dynamic analysis of urban crowd activities based on big data. Shanxi Normal University, Xi’an.
[3] Song L. (2017) Study on the accessibility of urban and rural public transport in Haining based on ArcGIS spatial network analysis. Chang’an University, Xi’an.
[4] Yaoli W., Song G. (2013) Exploration into urban street closeness centrality and its application methods: A case study of Qingdao, Geographical Research., 32(3): 452-464
[5] Yongqian L., Zhongzhen Y. (2018) Study on public transportation planning of small and mid-sized cities: Taking A county as an example, Journal of Guizhou University(Natural Sciences), 35(4): 122-124.