Research on Convenience Index of Urban Life Based on POI Data

Xinyang Ye, Hongli Tan, Yongzhi Zhang, Liang Zhang and Zhaohu Zhang

Digital Design Center, PowerChina Hubei Electric Engineering Corporation Limited, NO.1 Xinqiao 4th Road, Jinyinhu Street, Dongxihu District, Wuhan, China.

Email: Freda_yexinyang@163.com

Abstract. The urban life convenience index is the most important evaluation standard of the livable life. The convenience of residents' daily use of public and service facilities could represent the life convenience of people's urban life. Analytic Hierarchy Process (AHP) is a decision method that decomposes elements related to decision-making into levels such as objectives, criteria, and programs to perform qualitative and quantitative analysis. POI data has the advantages of high coverage of public service facilities and widely coverage of data, so it is effective to calculate the life convenience index and establish the system for assessing the convenience index of city life based on the POI data and AHP. It not only provides a practical reference for the special planning of urban public service facilities, but also explored good methods for planning applications using opened data.

1. Introduction

Livable cities have become widely accepted concepts of urban development. It could be said that the "livable city" is a comprehensive concept that involves the relationship between people and cities in terms of all living needs of people. The level of infrastructure and facilities that a city can provide is an important part of the evaluation of "Livability". In China, the evaluation criteria for livable cities are mainly based on the "China Livable City Evaluation Index System". As the basic and necessary condition for economic and social development, infrastructure is an energy for urban development, which could add stamina to urban development. Urban infrastructure construction relates to convenience of people’s lives and the urban life quality. The convenience of residents’ daily use of public and service facilities could represent the life convenience of people’s urban life. The urban life convenience index is the most important evaluation standard of the livable life. POI data is a point data that represents a real-world geographic entity. POIs generally contain basic information such as names, categories, latitude and longitude, and addresses. Some POIs obtained from specific websites for example, ratings, user reviews, and other data. POI data has the advantages of high coverage of public service facilities and widely coverage of data, so it is effective to calculate the life convenience index and establish the system for assessing the convenience index of city life based on the POI data and AHP. It not only provides a practical reference for the special planning of urban public service facilities, but also explored good methods for planning applications using opened data. The use of web crawler technology to obtain city POI data on Baidu maps, using POI data for research and analysis of topics is a basic way of combining big data with GIS. Major map service providers have opened the Application Programming Interface to facilitate developer access in order to allow more applications or websites to use map services more conveniently. For example, using Baidu map open platform, we could get the POI data and the Category.
Analytic Hierarchy Process (AHP) is a decision method that decomposes elements related to decision-making into levels such as objectives, criteria, and programs to perform qualitative and quantitative analysis. Analytic Hierarchy Process (AHP) categorizes various factors in a complex problem by dividing them into ordered levels that are related to each other, giving a quantitative representation of the relative importance of each level based on the judgment of a certain objective reality, and determining each expression using mathematical methods. The weight of the order of relative importance of all elements of a hierarchy.

Previous researches on life convenience index are mainly based on traditional methods such as questionnaires. The number of samples is small, the types are not comprehensive, and the distribution is uneven. It is difficult to accurately calculate the life convenience index from a macro perspective. With the improvement of people's living standards, the convenience of life has received more and more attention from the government and people. Therefore, it is necessary to find a suitable method to study the life convenience index. POI data can be a good expression of the number and spatial distribution of urban infrastructure. It is very effective to use the POI data to study the urban life convenience index. However, the sheer volume of POI data requires an effective way to integrate. At this time, the AHP can show the advantages of hierarchical expression of complex POI data.

2. Data Preparation
The study area in this paper is Guangzhou and Shenzhen. Guangzhou is the capital of Guangdong Province. Shenzhen is the first special economic zone established by China. Guangzhou and Shenzhen are all in Guangdong province and they are the most important cities in south of china. They have some similarities like the fast development of economy and urbanization or the large population. They are all cities with the most powerful economic strength and the most comprehensive strength in south of Guangdong province and also in the whole China, which are widely listed in the “first-tier cities”. This research intends to analyze and evaluate the convenience of residents living in these two cities based on the POI data of Guangzhou, and Shenzhen in 2018. Then, compared the advantages and disadvantages of the construction of these two cities. It mainly discusses the evaluation index system of city life convenience and the calculation method of the city life convenience index based on POI data, also trying to provide horizontal reference for the specific planning of public service facilities in these cities.

There are many ways to obtain POI data. The main method used in this paper is to use web crawler technology, write python language, and crawl data from Baidu map. There are 3 methods to crawl regional POI data. This study mainly used two methods: intra-city search and rectangular search. These two ways to achieve POI's regional search are both very convenient. However, due to the large size of the city, there is a large amount of POI information. Because of the restrictions Baidu's API, search within cities can only return up to 400 POI data. In order to prevent the phenomenon of the incomplete information. This research uses rectangular search to ensure that the number of POI data in each category does not exceed 400, so that the rectangles are searched one by one on a small grid, and all the POI data in the entire city area are finally obtained. In this study, a total of 1,239,368 POI data of Shenzhen were captured. There are 1,281,028 POI data of Guangzhou were obtained.

3. Analytic Hierarchy Process
The analytic hierarchy process can objectively describe people's subjective feelings, and quantitatively analyze non-quantitative events. The most critical of these is the quantitative quantification of subjective feelings of people. The project mainly studies the convenience of people’s urban life, and it is mainly aimed at the subjective feelings of residents. Therefore, the use of analytic hierarchy process is very appropriate.

3.1. Set up the Hierarchy Model
When using AHP to solve practical problems, we must organize the problem into hierarchical manner, and construct a hierarchical model with strong sense. In this study, a series of indicators used to assess the convenience of life were divided into two Hierarchy.
3.2. Build Judgment Matrix

After establishing a basic level of guidelines, we need to establish an judgment matrix with i-row and j-column based on the number of elements in the hierarchy. The criteria of each criterion level account for different proportions, and each of them occupy a certain proportion. When constructing the judgment matrix, the elements within the hierarchy are sorted according to the experience of many people. Sorting should take into account multiple factors, preferably using expert scoring. When the result of the sort is represented in a matrix, the inverse of 1-9 can be used as a scale to represent the matrix. Table 5 shows the definition of the judgment matrix scale.

**Table 1. The definition of the judgment matrix scale**

| Scale | meaning                                      |
|-------|----------------------------------------------|
| 1     | The two factors are equally important        |
| 3     | i is slightly more important than j than two factors |
| 5     | i is more important than j                   |
| 7     | i is significantly more important than j     |
| 9     | i is extremely important than j              |

Based on comprehensive considerations, the value of the matrix is in this order: Shopping service, Education service, Food service, Transportation services, Medical treatment facility, Convenient Service, Financial service, Leisure and entertainment. A 8*8 judgment matrix to represent the sorting of first level shows as follow.

\[
\begin{bmatrix}
1 & 5 & 3 & 3 & 5 & 5 & 7 & 7 \\
1/5 & 1 & 1/3 & 1/5 & 1 & 1 & 3 & 3 \\
1/3 & 3 & 1 & 3 & 5 & 5 & 7 & 7 \\
1/3 & 5 & 1/3 & 1 & 5 & 5 & 7 & 7 \\
1/5 & 1 & 1/5 & 1/5 & 1 & 1 & 3 & 3 \\
1/5 & 1 & 1/5 & 1/5 & 1 & 1 & 3 & 3 \\
1/7 & 1/3 & 1/7 & 1/7 & 1/3 & 1/3 & 1 & 1 \\
1/7 & 1/3 & 1/7 & 1/7 & 1/3 & 1/3 & 1 & 1 \\
\end{bmatrix}
\]

In the following analysis, I established eight second-level judgment matrices based on the surveys of people in Guangzhou and Shenzhen. Considering the needs of residents, the judgment matrices shows as below.

**Shopping service**

\[
\begin{bmatrix}
1 & 1/5 & 1/3 & 5 \\
5 & 1 & 3 & 9 \\
3 & 1/3 & 1 & 3 \\
1/5 & 1/9 & 1/3 & 1 \\
\end{bmatrix}
\]

**Food service**

\[
\begin{bmatrix}
1 & 3 & 5 & 5 \\
1/3 & 1 & 3 & 3 \\
1/5 & 1/3 & 1 & 1 \\
1/5 & 1/3 & 1 & 1 \\
\end{bmatrix}
\]

**Education service**

\[
\begin{bmatrix}
1 & 3 & 7 \\
1/3 & 1 & 3 \\
1/7 & 1/3 & 1 \\
\end{bmatrix}
\]

**Medical treatment facility**

\[
\begin{bmatrix}
1 & 3 & 3 \\
1/3 & 1 & 1 \\
1/3 & 1 & 1 \\
\end{bmatrix}
\]

**Transportation services**

\[
\begin{bmatrix}
1 & 1 \\
1 & 1 \\
\end{bmatrix}
\]

**Convenient Service**

\[
\begin{bmatrix}
1 & 3 \\
1/3 & 1 \\
\end{bmatrix}
\]

**Financial service**

\[
\begin{bmatrix}
1 & 3 & 5 \\
1/3 & 1 & 3 \\
1 & 1 \\
5 & 3 & 1 \\
\end{bmatrix}
\]

**Leisure and entertainment**

\[
\begin{bmatrix}
1 & 3 & 5 \\
1/3 & 1 & 3 \\
1 & 1 \\
5 & 3 & 1 \\
\end{bmatrix}
\]
3.3. Check the Consistency of the Single-order Hierarchical

After the judgment matrix is established, the consistency of the judgment matrix needs to be tested to judge the rationality of the matrix. The formula for the consistency ratio CR is:

\[ CR = \frac{CI}{RI} \]

When \( CR < 0.10 \), the judgment matrix is considered reasonable, otherwise it needs to be corrected. The formula for calculating the CI value is as follows:

\[ CI = \lambda_{\text{max}} \cdot \frac{n}{n-1} - 1 \]

\( \lambda_{\text{max}} \) is the maximum eigenvalue of the matrix, and \( n \) is the number of rows and columns of the matrix. The result of the calculation of the CR value is 0.0496, 0.0163, 0.0969, 0, 0.000175, 0, 0, 0.0370, 0.0370.

3.4. Calculate the Weight Vectors

Since each column in the judgment matrix approximately reflects the distribution of weights, the arithmetic average of all column vectors can be used to estimate the weight vector.

\[ W_i = \frac{a_{ij}}{\sum a_{kj}} / \sum \frac{a_{ij}}{\sum a_{kj}}, i = 1, 2, 3 ..., n \]

The value of the weights of each factor shows as the table 8 below.

**Table 2.** The weights of each factor

| level one          | Weight | level two                  | Condition weights | Normal weight |
|--------------------|--------|----------------------------|-------------------|---------------|
| Shopping service   | 0.3    | Convenience store          | 0.15              | 0.045         |
|                    |        | Department store           | 0.05              | 0.015         |
|                    |        | Supermarket chain          | 0.60              | 0.18          |
|                    |        | Agricultural Market        | 0.20              | 0.06          |
|                    |        | Kindergarten               | 0.70              | 0.07          |
| Education service  | 0.1    | primary school             | 0.20              | 0.02          |
|                    |        | Secondary school           | 0.10              | 0.01          |
|                    |        | Fast food                  | 0.60              | 0.12          |
| Food service       | 0.2    | Chinese restaurant         | 0.20              | 0.04          |
|                    |        | Foreign restaurants        | 0.10              | 0.02          |
|                    |        | Casual dining              | 0.10              | 0.02          |
|                    |        | Subway station             | 0.50              | 0.1           |
|                    |        | Bus stop                   | 0.50              | 0.1           |
| Transportation services | 0.2 | Specialist hospital        | 0.60              | 0.036         |
|                    |        | General Hospital           | 0.20              | 0.012         |
|                    |        | Pharmacy                   | 0.20              | 0.012         |
|                    |        | Beauty salons              | 0.75              | 0.045         |
|                    |        | Maintenance care           | 0.25              | 0.015         |
| Medical treatment facility | 0.06 | ATM machine                | 0.60              | 0.012         |
|                    |        | Four state-owned banks     | 0.30              | 0.006         |
|                    |        | Other banks                | 0.10              | 0.002         |
|                    |        | Children's playground      | 0.30              | 0.006         |
| Convenient Service | 0.06   | Theaters & stadiums        | 0.60              | 0.012         |
| Financial service  | 0.02   | Culture and art            | 0.10              | 0.002         |
| Leisure and        | 0.02   |                            |                   |               |
| entertainment      |        |                            |                   |               |
4. The Calculation of Life Convenience Index

4.1. Determine the Study Scale
This study combines the division of life circle theory and the observation of the distribution of communities in two cities. From the distribution, there are very few communities in the east of Shenzhen and in the north of Guangzhou. If the calculation of convenience is carried out within the scope of the entire city, errors may occur due to some certain special areas. Therefore, considering a variety of factors, I have determined a research criterion that the residential community should buffer the surrounding 500 m range (walking for 5-minute distance) and calculate the urban "community" life convenience index. 500-meter buffers centered on a residential area were shown.

4.2. Calculate the Distribution Density of Facilities
Using the tools of statistical analysis and area calculation in ArcGIS, the number of facilities in each buffer zone was obtained, and statistical analysis was performed to get the average values.

The Histogram of the Average number of facilities around the community shows as figure 1 below. According to the histogram, we can easily see the comparison of the Average number of facilities based on communities in both Shenzhen and Guangzhou. There are some differences between the two cities in the shopping service, food service and the convenience service. The analysis of this phenomenon would be explained in the follow part of this thesis.

![Average number of facilities around the community](image)

**Figure 1.** Average number of facilities around the community

4.3. Calculate the Life Convenience Index
After calculating the number of facilities statistically, the life Convenience index could be calculated. Since the determined research scale is performed on the established buffer zone, the area of each buffer zone is certain, so the average number of facilities can represent the density of the facility on each buffer zone.

For each evaluation criteria, which is represented by the average number of facilities, the score interval is normalized. Assume that the branch of each evaluation condition sees any value of Q in 0-1, and then use the highest value to normalize the score interval to obtain the Q value of each evaluation condition.

\[ Q_{ij} = \frac{N_{ij}}{\text{MAX}(N_{ij})} \]

\( N_{ij} \) represents the average number of facilities distributed in the jth facility in the i-th city.

The life convenience index could then be calculated by the formula below.

\[ Q_i = \sum Q_{ij} \times W \times W_c \]

\( Q_i \) represents the livability of urban living in the i-th city Number; W represents conditional weight; \( W_c \) represents the factor weight.

So the calculation result of the life convenience index shows as the table below.
### Table 3. Calculation of the life convenience index

| Level one         | Level two                  | Shenzhen | Guangzhou |
|------------------|----------------------------|----------|-----------|
| **Shopping service** | Convenience store | 0.023160905 | 0.045 |
|                  | Department store          | 0.00580895 | 0.015 |
|                  | Supermarket chain         | 0.18     | 0.161981614 |
|                  | Agricultural Market       | 0.000835655 | 0.06 |
| **Education service** | Kindergarten          | 0.07     | 0.05403183 |
|                  | primary school            | 0.008541667 | 0.02 |
|                  | Secondary school          | 0.005339806 | 0.01 |
|                  | Fast food                 | 0.113994225 | 0.12 |
| **Food service**  | Chinese restaurant        | 0.04     | 0.012094218 |
|                  | Foreign restaurants       | 0.02     | 0.014032496 |
|                  | Casual dining             | 0.002565267 | 0.02 |
|                  | Subway station            | 0.1      | 0.062162162 |
| **Transportation services** | Bus stop               | 0.094175824 | 0.1 |
|                  | Specialist hospital       | 0.031190661 | 0.036 |
|                  | General Hospital          | 0.012    | 0.004633205 |
| **Medical treatment facility** | Pharmacy              | 0.012    | 0.008758741 |
|                  | Beauty salons             | 0.045    | 0.0172319 |
| **Convenient Service** | Maintenance care        | 0.015    | 0.005420726 |
|                  | ATM machine               | 0.008258888 | 0.012 |
|                  | Four state-owned banks    | 0.00147541 | 0.006 |
| **Financial service** | Other banks               | 0.001468835 | 0.002 |
|                  | Children's playground     | 0.006    | 0.003095541 |
| **Leisure and entertainment** | Theaters & stadiums     | 0.006837209 | 0.012 |
|                  | Culture and art           | 0.001022472 | 0.002 |
| **Summary**      |                           | 0.804675773 | 0.803442435 |

#### 5. Conclusion

The above results show that, in general, there is no obvious difference in the convenience index of urban life in Guangzhou and Shenzhen. This may be due to the fact that the two cities belong to Guangdong Province and their living habits and routines are basically the same. For example, due to the warm climate in the south and relatively rich night-time activities, the corresponding public service facilities, such as dining and entertainment facilities, are relatively completed. Relatively speaking, the degree of community gathering is also relatively high. However, in spite of this, due to differences in cultural differences and population composition, Guangzhou and Shenzhen still have more obvious differences in some aspect.
Figure 2. The convenience index of each factors

With similar scores of living suitability scores, the scores of fast food restaurants near the Shenzhen community were significantly higher than those of Guangzhou, and the number of farmers' markets near the Guangzhou community was much greater than that of the Shenzhen community, indicating that the citizens of the two cities have different emphasis on suitability. The acceptance of dining out for Shenzhen residents is even higher, which may be due to the difference in the population structure between Shenzhen and Guangzhou; the cultural and educational facilities around the Guangzhou community are more complete than those in the Shenzhen community. Compared with Shenzhen, the cultural and educational facilities determine Guangzhou. The suitability index is an important factor. This is also due to the obvious difference in the culture of the two cities. In addition, the development of Shenzhen beauty salons may be due to the different demographic structure of the two cities. The proportion of young people in Shenzhen is relatively high. Therefore, the social focus of different demographics on the suitability index is different.

Shenzhen is a typical young immigrant city. Most of the residents are from all over the country. In addition, Shenzhen is a very young city, and most buildings are very new. In terms of cultural heritage and cultural development, there are obvious deficiencies in comparison with Guangzhou. According to statistics, the average age of resident population in Shenzhen is 32.5 years old, which is the youngest city in the country. This also determines that Shenzhen's development in entertainment facilities and beauty salons is slightly better than in Guangzhou. All in all, the suitability of the two cities is similar, but there are obvious differences in the factors that determine their suitability. This provides some guidance for the construction of livable cities. When evaluating and assessing livable cities, we should not only consider the urban infrastructure and objective living conditions, but also study the demographic structure and cultural atmosphere of each city. In this process, it is very difficult to establish a unified standard. At this time, the combination of research based on objective data (such as POI data) and research based on questionnaires is very necessary. POI data has a strong advantage in expressing spatial location and distribution density. This paper makes use of this advantage with taking into account the subjective factors such as urban population characteristics, which can provide a reliable reference for the special upgrade of urban infrastructure.

6. Reference
[1] Yuan J, Zheng Y, Xie X. Discovering regions of different functions in a city using human mobility and POIs[C]//Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2012: 186-194.
[2] Xie R, Zhang H, Gong C. Hot time periods discovery for facility proportioning in urban commercial districts using POIs and mobile phone data[C]//Web Intelligence. IOS Press, 2018, 16(2): 91-104.
[3] Vargas L G. An overview of the analytic hierarchy process and its applications[J]. European journal of operational research, 1990, 48(1): 2-8.