A comprehensive index model for real estate convenience based on multi-source data

Ling Hua*
School of Resource and Environmental Sciences, Wuhan University, Wuhan, China
*Corresponding author e-mail: 2018302050059@whu.edu.com

Abstract. The description information of the convenience of real estate is often qualitative and simple, and lacks comprehensive evaluation, so it is difficult to make horizontal comparison of the convenience of different real estate. Therefore, this paper proposes a comprehensive index model which uses multi-source data to measure the real estate convenience quantitatively. Based on the walking index, the model adds vector road network data, poi (point of interest) data and other multi-source data, and introduces the concept of walking score. Combined with the practical application of real estate, the model classifies the convenience facilities and calculates the walking distance between the convenience facilities and real estate quantitatively, so as to obtain a real number index to comprehensively measure the convenience of real estate. Finally, the experimental results based on Amap data prove the accuracy and effectiveness of the proposed model.

1. Introduction
Convenience of real estate refers to the degree of convenience of surrounding facilities and services and residents in real estate communities. It is an important content to measure the value of real estate and an important component of real estate advertising. It involves spatial information such as the spatial location of public facilities and the distance between them and real estate [1]. Therefore, as one of the big data, online map has been extended to the real estate field to describe the spatial distribution of real estate and its surrounding public facilities. This provides a wealth of information for people to understand the convenience of real estate, but these information lacks comprehensive and unified quantitative indicators.

In 2007, American researchers proposed the concept of "Walk score" based on daily facility layout, which mainly considered the types and spatial layout of daily facilities [2]. Meanwhile, factors such as walking distance attenuation, intersection density and block length were introduced to improve the accuracy of measurement. Walk score mainly reflects the rationality of daily facility allocation within a certain walking range, and has been determined as one of the bases to evaluate the efficiency of daily facility allocation in a city [3]. It can be used for the ranking of urban walkability, the real-time query of street walkability, the reference of rental price in a city and the research on the health level of urban community residents, which has been widely used in the United States, Canada, Australia and other countries [4]. Although walk score can reflect the convenience of real estate at a single point to a certain extent, it cannot be directly used to describe the convenience of real estate for two main reasons: first, the distance between a single point and the convenience facilities is a straight line, rather than the broken
path in reality; Second, the walk score excludes transportation facilities (such as subway entrance, bus station, etc.), while public transportation facilities directly affect the travel convenience of real estate [5].

This paper will build the comprehensive index model for real estate convenience based on the walk score model.

2. Preparation

2.1. Data sources and analysis

This paper mainly utilizes POI data and open source map data. The POI data comes from the background of Amap in 2020. Among them, eight types of facilities related to the convenience of real estate in Guangzhou are selected, including catering facilities, shopping services, transportation facilities, financial services, science and education, life services, sports leisure, and medical care. The open source map data comes from open source websites represented by OSM (Open Street Map). The road network data of Guangzhou is selected to obtain the information of road network longitude and latitude, street length and so on. The distribution of living facility points and road network in Guangzhou extracted from POI data and open source map data is shown in Figure 1.

![Figure 1. Distribution of living facilities and road network in Guangzhou.](image)

2.2. Walk score

The Walk Score is scored from 0 to 100 points, with higher scores indicating better walking. Walk score model mainly includes three parts [6]:

1. The types of public facilities. Walk score proposed by the United States divides amenities into grocery stores, restaurants and bars, shops, coffee shops, banks, parks, schools, bookstores, and entertainment venues [7]. According to domestic living habits, the types of convenience facilities were expanded, including schools, banks, and other new facilities;

2. The Euclidean straight-line distance between the facilities and the center;

3. Distance attenuation coefficient of facility convenience. Walk Score assigns different attenuation intensities to facilities distributed in different distances, and expands the ratio to 0 to 100 to get the final walk score.

3. Real estate convenience index model

The real estate convenience index is a comprehensive index model to measure the convenience of real estate, and an extension of the walk score model in the real estate field.

Like the walk score model, the convenience index model of real estate is also divided into three parts: facility type, distance between center point and facility, and attenuation coefficient of facility convenience.
(1) Facilities type of convenience index. Some real estate websites differ in their classification of amenities surrounding real estate. According to gz.fang.com, it is divided into transportation, education, life, health, and catering; According to gz.julive.com, it is divided into transportation, education, business, hospital, etc. Although the classification of supporting facilities around real estate is different, it is generally consistent. In order to express the convenience of real estate in a more detailed way, the classification of surrounding supporting facilities should cover the daily travel needs of residents as comprehensively as possible. The classification table of facilities constructed in this paper includes 28 types of facilities in 8 aspects, including catering facilities, shopping services, transportation facilities, financial services, science and education, life services, sports leisure, and medical care [8]. Each type of facilities is given a weight according to its relative importance (as shown in table 1), and the weight sum is 1. (The investigation of the initial weight value is beyond the scope of this paper. It is worth emphasizing that the uniform weighting criteria can distinguish between different properties in terms of convenience.)

Define the weight of first-level index \( i \) as \( W_i \), where \( i = 1, 2, 3... \) Define the weight of the second-level index \( j \) as \( W_{i,j} \), and the sum of the weights of the second-level index is equal to the weight of the first-level index.

\[
\sum_j W_{i,j} = W_i \quad (1)
\]

The weight sum of all first-level indicators and second-level indicators equals 1,

\[
\sum_i W_i = \sum_i \sum_j W_{i,j} = 1 \quad (2)
\]

**Table 1. Weights Distribution of Real Estate Ancillary Facilities.**

| First-grade Indexes       | Weights of first-level indicators | second-grade indexes                          | Weights of second-level indicators |
|---------------------------|----------------------------------|-----------------------------------------------|-----------------------------------|
| Catering Facilities       | 0.1957                           |                                  | Chinese food restaurant 0.100     |
|                           |                                  |                                  | Casual restaurant 0.031          |
|                           |                                  |                                  | Foreign restaurant 0.0146        |
|                           |                                  |                                  | Fast-food restaurant 0.0501      |
|                           |                                  |                                  | Convenience store 0.0389         |
| Shopping Service          | 0.3235                           |                                  | Supermarket 0.0858               |
|                           |                                  |                                  | Shopping centers, shopping malls 0.1988 |
| Transportation Facilities | 0.1204                           |                                  | Subway station 0.0574            |
|                           |                                  |                                  | Bus stop 0.0473                  |
|                           |                                  |                                  | Parking 0.0157                   |
| Financial Service         | 0.0378                           |                                  | ATM 0.0237                       |
|                           |                                  |                                  | Bank 0.0106                      |
|                           |                                  |                                  | Finance company 0.0035           |
| Science and education     | 0.0700                           |                                  | Kindergarten 0.0346              |
|                           |                                  |                                  | Primary school 0.0189            |
|                           |                                  |                                  | Middle school 0.0108             |
|                           |                                  |                                  | Training institutions 0.0057     |
| Life Service              | 0.1163                           |                                  | Home-moving company 0.0075       |
|                           |                                  |                                  | Telecommunications business hall 0.0613 |
|                           |                                  |                                  | Beauty and hair dressing 0.0324  |
| Sports Leisure            | 0.0536                           |                                  | Maintenance point 0.0153         |
|                           |                                  |                                  | Sports venues 0.0063             |
|                           |                                  |                                  | Entertainment venue 0.0144       |
|                           |                                  |                                  | Leisure space 0.0329             |
| Medical Care              | 0.0827                           |                                  | Comprehensive hospitals 0.0291   |
|                           |                                  |                                  | Specialized hospitals 0.0200     |
|                           |                                  |                                  | Clinic 0.0206                    |
|                           |                                  |                                  | Pharmacy 0.013                   |
The distance measure of the convenience index. In the description of real estate convenience, the distance between public facilities and real estate is an important content. The walk score model uses linear distance to measure the spatial distance between the central location and public facilities, ignoring the influence of road network and community environment. In reality, the path between two points is usually not a straight line, but a broken line along a block road. Therefore, the influence of the actual environment is further considered, and the straight-line distance is corrected through information such as intersection density and block length measurement. The attenuation rate of intersection density and block length is divided into five levels (as shown in Table 2).

Table 2. Comparison table of intersection density and block length decay rate.

| Intersection density (/km²) | Decay_rate1(%) | Block length(m/100) | Decay_rate2(%) |
|----------------------------|----------------|---------------------|----------------|
| ≥200                       | 0              | ≤120                | 0              |
| 150-200                    | 1              | 120-150             | 1              |
| 120-150                    | 2              | 150-165             | 2              |
| 90-120                     | 3              | 165-180             | 3              |
| 60-90                      | 4              | 180-195             | 4              |
| <60                        | 5              | >195                | 5              |

Decay coefficient of convenience index. The convenience of public facilities is distance-attenuating. Based on the classification table of facilities, considering the distance between public facilities and real estate, the initial weight of facilities will decrease regularly with the increase of the distance from the starting point. In this paper, cubic curve is adopted as the distance decay rule. According to the standard of distance decay rule of Walkscore.com and the standard walking speed of 80 m/min, the following distance attenuation law is obtained, as shown in Table 3 (the distance between the centre point and public facilities is \( d \)).

Table 3. Distance Decay Function of Facilities Convenience.

| Time   | Scope of Arrival | Distance Decay Rule | Remark                           |
|--------|------------------|---------------------|---------------------------------|
| -5min  | 400m             | No distance attenuation, \( f(d)=1 \). | -                 |
| -20min | 1600m            | Rapid decay: When it is located at 1600 m, the attenuation rate attenuates to 12%. \( f(d) = -153.6558d^3 + 419.4604d^2 - 395.9706d + 201.1086 \) | -                 |
| -30min | 2400m            | The attenuation slows down, and when the distance is greater than 2400m, the attenuation rate is greater than 1. \( f(d) = -92.8d^3 + 566.6d^2 - 1153.1d + 786.6 \). Facilities outside 2400m have no effect on the starting point's walk index. | -                 |

The convenience index model of any central point is shown in equation (3).

\[
RECS = \sum_i \sum_j (W_{i,j} \times f(d)) \times (100 - \text{decay}_\text{rate1}) \times (100 - \text{decay}_\text{rate2}) / 100 \tag{3}
\]

4. Experiment and result analysis

4.1. Real Estate Convenience Calculation

The experimental object of this paper is the real estate in Guangzhou, which are recorded as Estate #1 and Estate #2 respectively. According to the data sources of gz.julive.com and Amap, combined with the facility classification of convenience index model, a classification and grading data of real estate and surrounding supporting facilities and the corresponding path distance are obtained. The attenuation rates of intersection density of Estate #1 and Estate #2 are both 0.05, and the attenuation rates of block length are both 0. According to Formula 3, the obtained results are normalized, that is, the convenience index with a value of 0-100 is obtained. In the end, the convenience index of Estate #1 was 73.60(The value before normalization:325.1), as shown in table 4. Similarly, it is possible to obtain Estate #2 with a convenience score of 96.31 (The value before normalization:423.5).
Table 4. Convenience score of Estate #1 in Guangzhou.

| First-grade Indexes       | Score | Second-grade indexes       | Score |
|---------------------------|-------|---------------------------|-------|
| Catering Facilities       | 0.6551| Chinese food restaurant   | 0.5196|
|                           |       | Casual restaurant         | 0.0029|
|                           |       | Foreign restaurant        | 0.0131|
|                           |       | Fast-food restaurant      | 0.1195|
| Shopping Service          | 0.9620| Convenience store         | 0.0938|
|                           |       | Supermarket               | 0.6998|
|                           |       | Shopping centers, shopping malls | 0.1684|
| Transportation Facilities | 0.8038| Subway station            | 0.2319|
|                           |       | Bus stop                  | 0.3802|
|                           |       | Parking                   | 0.1917|
|                           |       | ATM                       | 0.0283|
|                           |       | Bank                      | 0.0212|
| Financial Service         | 0.0568| Finance company           | 0.0073|
|                           |       | Kindergarten              | 0.0528|
|                           |       | Primary school            | 0.0758|
|                           |       | Middle school             | 0.0229|
|                           |       | Training institutions     | 0.0270|
| Science and education     | 0.1785| Home-moving company       | 0.0013|
|                           |       | Telecommunications business Hall | 0.0105|
|                           |       | Beauty and hair dressing  | 0.1777|
|                           |       | Maintenance point         | 0.0437|
| Sports Leisure            | 0.1870| Sports venues             | 0.0732|
|                           |       | Entertainment venue       | 0.0847|
|                           |       | Leisure space             | 0.0291|
| Medical Care              | 0.1762| Comprehensive hospitals    | 0.0431|
|                           |       | Specialized hospitals     | 0.0450|
|                           |       | Clinic                    | 0.0340|
|                           |       | Pharmacy                  | 0.0551|

4.2. Real Estate Convenience Verification

Estate #1 and Estate #2 are both located in the central urban area of Guangzhou, with a complete range of surrounding facilities. Therefore, the convenience index of the two properties is relatively high. At the same time, due to the large distance difference between the two facilities, the convenience is not the same. The quantitative reflection of this difference is the convenience index of the poor.

The price difference between Estate #1 and Estate #2 also reflects the convenience difference between the two. According to gz.fang.com, gz.julive.com, etc., the average price of Estate #1 is about 17,000 yuan/m², while the average price of Estate #2 is about 23,000 yuan/m², and the house price of Kanazawa Court is slightly higher than Estate #1. Estate #1 and Estate #2 are in a similar location, with a distance of about 3km between them. It is the convenience of surrounding facilities that makes the difference between the two housing prices. Therefore, the convenience index based on the convenience of the surrounding facilities of real estate can effectively support the horizontal comparison of the convenience of two real estates, so as to provide a simple and unified mathematical basis for people to understand the convenience difference between real estate.

5. Conclusions

In this paper, by collecting road network vector data and POI data, based on the measurement method of walk score and combining with the actual application of real estate, a comprehensive index model of real estate convenience is constructed from eight dimensions. The convenience index model is used to measure the convenience of two real estate locations to verify the effectiveness of the algorithm. The comprehensive index method of real estate convenience combines the traditional planning idea with the new data environment, and then provides more scientific and rigorous data support. It has the characteristics of comprehensive, quantification and objectivity, and provides a new idea for the quantification and thematic research of the surrounding facilities information of urban real estate.
References

[1] Shanna Trichès Lucchesi, et al. Are people willing to pay more to live in a walking environment? A multigroup analysis of the impact of walkability on real estate values and their moderation effects in two Global South cities. Research in Transportation Economics. 2020, in press.

[2] Hao Wang, and Yuqi Yang. Neighbourhood walkability: A review and bibliometric analysis. Cities. 2019, 93:43-61.

[3] Carr, Lucas & Dunsiger, Shira & Marcus, Bess. (2010). Walk Score (TM) As a Global Estimate of Neighborhood Walkability. American journal of preventive medicine. 39. 460-3. 10.1016/j.amepre.2010.07.007

[4] Jianxiao Liu, BI Han, and Meilian Wang. Using multi-source data to assess livability in Hong Kong at the community-based level: A combined subjective-objective approach. Geography and Sustainability. 2020, 1(4): 284-294.

[5] Mohammed Ali Berawi, et al. Impact of rail transit station proximity to commercial property prices: utilizing big data in urban real estate. Journal of Big Data. 2020, 7(1): 71

[6] Girling, Zheng, Monti, et al. Walkability vs. walking: assessing outcomes of walkability at Southeast False Creek, Vancouver, Canada. 2019, 12(4):456-475.

[7] Z. Yazicioglu Halu. Transactional Approach for Walkable Urban Spaces: Hierarchy of Walking Needs. Journal of Environmental Protection and Ecology. 2019, 20(1):302-312.

[8] Teng Zhong, Guonian Lü, Xiuming Zhong, Haoming Tang, Yu Ye. Measuring Human-Scale Living Convenience through Multi-Sourced Urban Data and a Geodesign Approach: Buildings as Analytical Units. Sustainability. 2020, 12(11): 4712.