

# Urban Functional Area Identification based on POI Data

Yajing Zhou \*

Department of Economics and Management, Hebei University of Technology, Tianjin, China

\* Corresponding author Email: 2625350755@qq.com

---

**Abstract:** With the deepening of urbanization, the urbanization process is gradually accelerated, and the urban spatial structure is constantly changing. The identification of urban functional areas is of great significance for optimizing the urban spatial structure and analyzing the behavior characteristics of residents. The emergence of geographic information data such as POI provides a new perspective for the study of urban functional area identification. This paper illustrates the identification of urban functional area based on POI from two perspectives: the data sources commonly used in the field of urban functional area identification and the methods commonly used in the field of urban functional area identification.

**Keywords:** POI Data; Urban Functional; Machine Learning.

---

## 1. Introduction

The identification of urban functional areas refers to the division of urban spatial areas according to urban spatial functions to help people deepen their understanding of urban physical space and social space. As an important part of urban economic and social functions, urban functional areas meet the daily needs of urban residents for work, residence and entertainment. Data commonly used in urban functional area identification include remote sensing image data, questionnaire survey data, point of interest (POI) data, social media data, mobile phone data, public transport card swipe data and taxi track data. Common identification methods of urban functional areas include qualitative analysis, cluster analysis, machine learning and so on.

## 2. Research Status of Urban Functional Area Identification

At the Fourth International Congress of Modern Architecture held in 1993, scholars put forward the concept of "urban functional district" for the first time according to the relationship between architecture and urban planning. Urban functional areas are defined as areas with similar land use and land use potential [1] in which the same type of natural resources and social services tend to be concentrated. As an important part of urban economic and social functions, different types of urban functional areas meet the daily work, living, entertainment and other needs of urban residents. Common urban functional areas include office areas, residential areas, leisure areas, etc.

Urban functional zone identification is often applied to urban structure optimization, urban resource allocation, user behavior analysis and other fields [2]. For example, Yang et al [3] divided Xi 'an into 422 research areas based on administrative boundaries, and built a comprehensive urban travel grid based on GPS track data to provide suggestions for urban operation management and urban spatial layout. Wang et al [4] used the road network to divide the basic units of urban functional area identification, and based on taxi track data and point of interest data, proposed a new research method of multi-centrality of urban functions to help urban authorities better understand the urban dynamics in terms of functional distribution and internal connectivity, so as to

provide references for urban authorities to allocate resources. Yang Zhenshan et al [5] divided Beijing into a regular grid of 250 m×250 m. Based on cell phone signaling data and point of interest data, they quantitatively studied the day and night difference of functional use intensity and the degree of internal functional mixing within the region, and found that service facilities such as catering and living facilities had a higher intensity of use at night, while facilities such as financial and public services had a higher intensity of use during the day.

Traditional data used in urban functional area identification include remote sensing image data and questionnaire survey data. However, remote sensing image data reflects urban land cover through physical features such as spectrum and texture of ground objects in remote sensing images, but does not reflect social and economic conditions. However, questionnaire survey not only consumes a lot of manpower, but also contains subjective factors of investigators in the data obtained from questionnaire survey. With the rapid development of communication technology, massive spatio-temporal data has been generated in cities, including point of interest (POI) data, social media data, mobile phone data, public transport card data and taxi track data, etc. These spatio-temporal data have complete content and low acquisition cost. It brings new opportunities for the study of urban spatial structure. Social media data contains rich semantic information, but it is difficult to extract the useful semantic information for functional area recognition. Mobile phone data can truly reflect the real-time location information of residents, but there is no information about the city function; Public transport card swipe data and taxi track data more reflect the travel behavior of residents; POI data clearly defines activity types and locations, and is closely related to land use types. Therefore, POI data can be used as spatial semantic annotation functional areas to reflect social functional attributes of different urban land. POI data has been widely used in the field of identification of urban functional areas. For example, Zhen [6] has identified urban functional areas in the central city of Changchun through quantifying POI data, and compared the identification results of functional areas with the current land use map of Changchun, verifying the accuracy of the proposed method of identification of functional areas. Cao et al [7] calculated the weighted frequency-density ratio of each POI inside a

building and automatically classified a large number of buildings into different functional types, improving the quantitative identification method of urban functional areas, and the classification method based on density analysis is easy to understand. Chen et al [8] proposed a new idea of using POI data to identify urban functional areas, that is, to understand the spatial organization structure of urban functional areas by mining the co-occurrence pattern of POI data in cities, and to identify the functionality and characteristics of the organizational structure of urban functional areas by taking 25 domestic cities as research objects. They found that although the structure of cities is relatively similar, However, the co-occurrence patterns of POI data in different functional areas are quite different.

In the identification of urban functional area, the research area is usually divided into several spatial units as the basic unit of urban functional area identification. The division method and scale of spatial units directly affect the identification accuracy of urban functional area. The common division of space units includes three ways: administrative area [3], road network [4] and grid [5, 9-10]. The division of spatial units based on administrative boundaries refers to taking the original administrative divisions of the city such as streets and communities as the smallest units. But in practice, the administrative boundary becomes blurred in the complex urban system, which makes the division of space units difficult. The division of spatial units based on road network data is to divide the research area according to road network data in the city, and take the specific area surrounded by roads as the basic research unit. However, due to the irregular distribution of the road network, the identified spatial units are not closed, so it is necessary to conduct secondary processing of the spatial units in the follow-up study. However, the most common method of spatial cell division is mesh-based spatial cell division, that is, the research area is divided into equal size grids. The grid - based division of space cells not only simplify the problem model, but also reduces the computational complexity, and is easy to operate and understand. For example, Liu Ju et al [9] divided the Sixth Ring Road of Beijing into a 500 m×500 m grid, constructed the O point and D point tensors, and used the tensor decomposition model to reveal the travel time pattern of taxi users from the daily scale and time period scale. Wang et al [10] divided Beijing into a 1 km×1 km regular grid, and proposed a functional area identification method based on spatial semantics based on taxi OD point data and POI data. Firstly, the research area was divided into three functional areas by using taxi OD point data, and then the spatial semantics of the three functional areas were analyzed by using point of interest data. The intensity of interaction between functional areas was quantitatively discussed from the perspective of temporal and spatial analysis. It is worth noting that the grid is the smallest unit for urban functional area identification, so it is necessary to pay attention to the accuracy of the grid in the process of urban functional area identification, and determine the grid specifications according to different research needs and accuracy.

Common methods of urban functional area identification

### 3. Common Methods of Urban Functional Area Identification

Common identification methods of urban functional areas include qualitative analysis, cluster analysis, machine

learning and so on. Qualitative analysis includes statistical survey, expert voting and other methods, most of which realize fast and efficient classification of urban functional areas through GIS platform. For example, Hu and Han [11] took Guangzhou Economic and Technological Development Zone as an example and proposed an urban functional area identification method based on frequency density and POI type ratio based on POI data to analyze the main functions and spatial distribution characteristics of urban functional areas in detail. Although qualitative analysis method can realize the division of urban functional areas based on public will, it is time-consuming and laborious, with strong subjectivity, and difficult to prove the accuracy of the results, so the accuracy needs to be improved [12].

Cluster analysis is to discover the spatial and temporal characteristics of residents' travel based on static or dynamic urban data, such as mobile phone data, POI data and taxi track data, and identify urban functional areas according to the spatial and temporal characteristics of residents' travel in different areas. Common clustering algorithms include density-based clustering, hierarchy-based clustering and partition-based clustering, among which the representative ones are DBSCAN, OPTICS and K-means. For example, Chen Zhanlong et al [13] used the spatial co-location pattern to dig out the distribution features of POI and form the regional feature vector, and used the K-means clustering algorithm for regional clustering to identify different types of urban functional areas in Beijing, such as entertainment business district, science, education and culture district, etc. Li Ke and Dang Yanzhong [14] adopted the density-based peak clustering algorithm CFSFDP to cluster the data of taxi pick-up and pick-up points. After discovering the hot space areas of residents' travel, they identified different types of urban functional areas through expert identification method. Cluster analysis has good accuracy for urban function identification, but in complex urban areas, clustering accuracy will be affected to some extent.

The machine learning method is based on the relationship between data, and uses the machine learning model to extract features and identify urban functional areas. Machine learning algorithm is suitable for extracting high-level information, and its portability and scalability are very suitable for identifying and analyzing urban functional areas. For example, Sun Shijie et al [15] used random forest algorithm to evaluate the importance of features based on the number of taxi passengers on working days and rest days in different regions as the feature of machine learning, and finally adopted classical decision tree algorithm to classify functional areas. Taking Yuzhong District of Chongqing as an example, Cao et al [16] compared XGBoost algorithm with polynomial Logistic regression, k-nearest neighbor, decision tree, support vector machine (SVM), random forest and other machine learning classification algorithms based on night images, geo-tagged microblog check-in data, POI data and other multi-source data. It is found that the accuracy rate of XGBoost algorithm can reach 88.05%, which shows that XGBoost algorithm has achieved a good classification effect in the identification of urban functional areas.

In addition, deep learning algorithms such as convolutional neural networks (CNNs) have been widely used in urban functional area recognition in recent years due to their powerful image processing and feature extraction capabilities. For example, Cao et al [17] proposed a new remote sensing and social sensing data fusion model based on end-to-end

deep learning, which automatically extracts time-related social perception features, integrates these features with remote sensing image features extracted through residual neural network, and conducts a large number of experiments on real data sets. The proposed method is proved to be efficient in the field of urban functional area identification. Hu et al [18] introduced a framework of geographic semantic analysis to explore the interaction mode between urban functional structure and traffic space by modeling the traffic interaction between road segments and the relationship between cities. First, the road trajectory corpus was constructed and trained to obtain the semantic embedding representation of road segments. The convolutional neural network model was used to process the context and topological information, and the social functions along the street were classified. Finally, the model was verified based on the rental trajectory of Beijing.

## References

- [1] Gao S, Janowicz K, Couclelis H J T G. Extracting urban functional regions from points of interest and human activities on location-based social networks[J]. *Transaction in GIS*, 2017, 21(3): 446-467.
- [2] Liu B, Deng Y, Li M, et al. Classification schemes and identification methods for urban functional zone: A review of recent papers[J]. *Applied Sciences*, 2021, 11(21): 9968.
- [3] Yang Y, He Z Z, Song Z, et al. Investigation on structural and spatial characteristics of taxi trip trajectory network in Xi'an, China[J]. *Physica A: Statistical Mechanics its Applications*, 2018, 506(15): 755-766.
- [4] Wang S, Yanbin L, Zhi W, et al. Discovering urban functional polycentricity: A traffic flow-embedded and topic modeling-based methodology framework[J]. *Sustainability*, 2020, 12 (5): 1857.
- [5] Yang Z S, Su J H, Yang H, et al. Exploring urban functional areas based on multi-source data: A case study of Beijing [J]. *Geographical Research*, 2021, 40(02): 477-494.
- [6] Zhen Q. Quantitative Identification of urban functional areas in downtown area of Changchun based on POI data[J]. *IOP Conference Series: Earth Environmental Science*, 2019, 330(5): 052001.
- [7] Cao Y H, Liu J P, Wang Y, et al. A Study on the method for functional classification of urban buildings by using POI data [J]. *Journal of Geo-Information Science*, 2020, 22(6): 1339-1348.
- [8] Chen Y, Chen X, Liu Z F, et al. Understanding the spatial organization of urban functions based on co-location patterns mining: A comparative analysis for 25 Chinese cities[J]. *Cities*, 2020, 97: 102563.
- [9] Liu J, Xu J, CAI Ling, et al. Identifying Functional Regions Based on the Spatio-temporal Pattern of Taxi Trajectories[J]. *Journal of Geo-Information Science*, 2018, 20(11): 1550-1561.
- [10] Wang Y, Gu Y, Dou M, et al. Using spatial semantics and interactions to identify urban functional regions[J]. *ISPRS International Journal of Geo-Information*, 2018, 7(4): 130.
- [11] Hu Y, Han Y K. Identification of urban functional areas based on POI data: A case study of the Guangzhou economic and technological development zone[J]. *Sustainability*, 2019, 11 (5): 1385.
- [12] Yuan G, Chen Y, Sun L, et al. Recognition of functional areas based on call detail records and point of interest data[J]. *Journal of Advanced Transportation*, 2020: 1-16. DOI: 10.1155/ 2020/ 8956910.
- [13] Chen Z L, Zhou L L, Yu W H, et al. Identification of the urban functional regions considering the potential context of interest points [J]. *Journal of Geodesy and Geo-information Science*, 2020, 49(07): 907-920.
- [14] Li K, Dang Y Z. The Precise Delivery Strategy of LED Advertisement in New Operational Mode of Taxi [J]. *Journal of Management Science and Engineering*, 2020, 192(10): 220-230.
- [15] Sun S J, Sun Q, Lu C W, et al. Urban functional Areas Extraction Method Based on Taxi Disembarkation data [J]. *Journal of Science and Technology of Surveying and Mapping*, 2019, 36(06): 637-642.
- [16] Cao K, Guo H, Zhang Y. Comparison of approaches for urban functional zones classification based on multi-source geo-spatial data: A case study in Yuzhong District, Chongqing, China [J]. *Sustainability*, 2019, 11(3): 660.
- [17] Cao R, Tu W H, Yang C, et al. Deep learning-based remote and social sensing data fusion for urban region function recognition [J]. *Isprs Journal of Photogrammetry Remote Sensing*, 2020, 163: 82-97.
- [18] Hu S, Gao S, Wu L, et al. Urban function classification at road segment level using taxi trajectory data: A graph convolutional neural network approach[J]. *Computers Environment and Urban Systems*, 2021, 87(8): 101619.