

# An Automatic Urban Function District Division Method Based on Big Data Analysis of POI

Hao Guo\*, Haiqing Liu\*\*, Shengli Wang\*\*\*, and Yu Zhang\*\*

## Abstract

Along with the rapid development of the economy, the urban scale has extended rapidly, leading to the formation of different types of urban function districts (UFDs), such as central business, residential and industrial districts. Recognizing the spatial distributions of these districts is of great significance to manage the evolving role of urban planning and further help in developing reliable urban planning programs. In this paper, we propose an automatic UFD division method based on big data analysis of point of interest (POI) data. Considering that the distribution of POI data is unbalanced in a geographic space, a dichotomy-based data retrieval method was used to improve the efficiency of the data crawling process. Further, a POI spatial feature analysis method based on the mean shift algorithm is proposed, where data points with similar attributive characteristics are clustered to form the function districts. The proposed method was thoroughly tested in an actual urban case scenario and the results show its superior performance. Further, the suitability of fit to practical situations reaches 88.4%, demonstrating a reasonable UFD division result.

## Keywords

Big Data Analysis, Dichotomy Method, Mean Shift Algorithm, POI Data, Urban Function District

## 1. Introduction

With the accelerating urbanization process, the city size has extended, inducing several problems, such as increased traffic congestion and pollution. To solve these problems, reasonable urban planning should be considered before undertaking urban development. The urban function district (UFD) is the most important concept in urban planning for describing urban structures. In detail, the UFD indicates a certain geographic area where the land use function, intensity, direction, and price are consistent, and the intensive use potential also has similar features. UFD division is highly significant to handle the evolving role of urban planning and further help in developing reliable urban planning programs [1,2].

Conventionally, UFD division is mainly based on investigative methods, such as spot surveys or questionnaires [3]. In recent years, newly developed technologies, such as remote sensing [4,5] and unmanned aerial vehicle (UAV) photography [6,7] are also being used for the task. However, these methods require vast amounts of survey data and consume significant manpower and time. In addition, the division results are affected by the personal experience of the planner and subjective factors. Furthermore, with the accelerating urbanization process, it is also difficult for planners to find newly

※ This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Manuscript received July 4, 2019; first revision June 2, 2020; accepted June 9, 2020.

**Corresponding Author:** Haiqing Liu (hqliu@sdu.edu.cn)

\* College of Geomatics, Shandong University of Science and Technology, Qingdao, China (guoritian94@gmail.com)

\*\* College of Transportation, Shandong University of Science and Technology, Qingdao, China (hqliu@sdu.edu.cn, zy-668@qq.com)

\*\*\* Ocean Science and Engineering College, Shandong University of Science and Technology, Qingdao, China (shlwang@sdu.edu.cn)

formed function districts without obvious characteristics, rendering the traditional investigation methods unsuitable.

As a worldwide technology trend, big data analysis has been widely used in several fields and involves the discovery of useful implicit information in large datasets using data extraction, machine learning, statistics, and visualization. Compared with the traditional methods, big data analysis can provide a highly objective, reasonable, and visualized UFD division result. In recent years, several scholars have applied this new technology to UFD. Based on the literature review, recent UFD division methods based on big data analysis can be classified into three types: (1) methods based on cellular signaling data, (2) methods based on remote sensing images and point of interest (POI) data, and (3) methods based on vehicle trajectory, such as the probe data of transit vehicles and taxis.

Researchers have achieved reasonable success with these methods in recent years. For example, using cellular signaling data, Yan et al. [8] presented an urban functional area division method based on GIS clustering analysis. The method was verified using actual data from Changchun City, China, and the results showed excellent performance. Zhang et al. [9] presented a hierarchical semantic cognition (HSC) method to classify functional zones in Beijing. The proposed method relies on geographic cognition and considers visual features, object categories, spatial object patterns, zone functions, and hierarchical relations. It can produce a high overall accuracy of 90.8%. In [10], the authors divided the urban function area with a DBSCAN clustering method using the trajectory data of 4,000 taxis. In this method, the passengers' pick-up and drop-off states are extracted from the trajectory for the division, and the total accuracy can reach 95%.

In the aforementioned works, the most commonly used methods for big data analysis are clustering, such as K-means and K-medoids. These algorithms belong to supervised classification approaches in which certain clustering centers must be presumed. Given different clustering centers, the division results are different [11]. In urban scenarios, UFDs may change with the growing urbanization, leading to generally known clustering centers, especially in a few new city areas. In addition, certain data resources are difficult to acquire and process, such as remote sensing images. Cellular signaling data involves individual privacy, and is not open to the public, considering the legal viewpoint. These shortcomings render the traditional methods inadequate for achieving high-efficiency, low-cost, and dynamic UFD division.

In this paper, an automatic UFD division method is proposed based on POI data analysis. To improve the data retrieval efficiency in case of an unbalanced distribution of data points in a geographic space, the dichotomy method was used to optimize the data extraction process by crawling on the Internet. Based on this, the UFD was divided using the mean shift method. The overall work was tested in an actual urban scenario, and the results demonstrate the credible performance of the proposed method.

The remainder of this paper is organized as follows. In Section 2, the dichotomy-based POI data extraction method is presented. In Section 3, the UFD division method using the mean shift algorithm is proposed. The case study, along with the data extraction performance and UFD division results, are presented in Section 4. In Section 5, we conclude the paper.

## 2. POI Data Extraction Method based on Dichotomy

### 2.1 Introduction of the POI Data

POI implies a specific point location that may be considered useful or interesting by people, such as a

store, hotel, campsite, fuel station, or toponym with humanistic significance [12]. The positional information for the POI data is acquired using precision positioning equipment or map orienting. The data contain vast quantities of spatial, attributive, and other useful information, such as comments about the shop and travel photos. This information provides new data support for automatic UFD division [13]. In this study, the POI data were completely acquired using a web crawler. The acquired POI data were divided into 10 categories and 68 sub-categories. The overall data distribution and categories of the POI data are presented in Fig. 1 and Table 1, respectively.

Each data item represents the name, longitude, latitude, administrative region, and other information, as shown in Table 2.

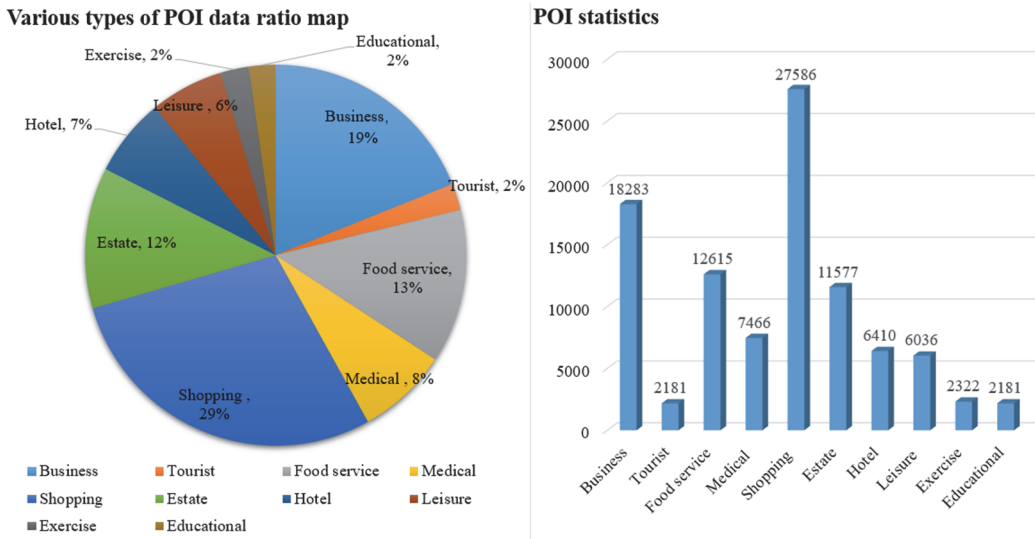


Fig. 1. Proportion of various types of POI data.

Table 1. Categories of POI data

Category	Sub-category
Food service	Chinese restaurant, foreign restaurant, snack bar, cake shop, cafe, tea house, bar
Shopping service	Shopping center, department stores, supermarkets, convenience stores, home building materials, home appliances, digital stores, markets
Hotel	Star hotel, express hotel, apartment hotel
Tourist attractions	Parks, zoos, botanical gardens, amusement parks, museums, aquariums, beaches, monuments, churches, scenic spots
Real estate	Office building, residential area, dormitory
Medical service	General hospitals, specialized hospitals, clinics, pharmacies, medical institutions, nursing homes, first aid, disease control center
Leisure and entertainment	Holiday village, farmyard, cinema, KTV, theatre, dance hall, Internet bar, gaming place, bath and massage, leisure square
Exercise and fitness	Sports venues, extreme sports venues, fitness centers
Incorporated business	Companies, parks, agriculture, forestry and horticulture, factories and mines
Educational training	Institutions of higher learning, secondary schools, primary schools, kindergartens, adult education, parent-child education, special education schools, overseas study agencies, scientific research institutions, training institutions, libraries, science and technology museums

**Table 2.** The description for each field in the case POI data

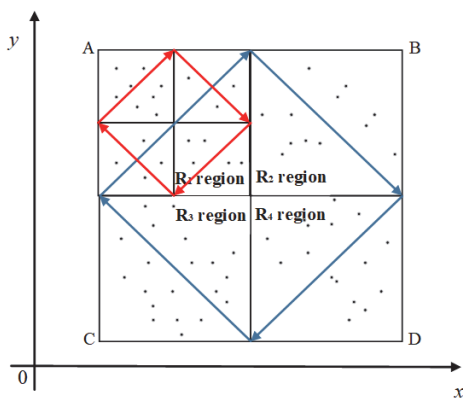
Field	Description
Name	Data name
LAT	Latitude value
LONG	Longitude value
Address	Detailed address information
Province	Province information
City	City name
Area	Address of the area
Street_ID	ID of the street
Uid	Unique identity
Categories	Ten broad categories of numerated assignment classification
Classification	Sub-classifications of the broad categories

### 2.2 POI Big Data Extraction based on Dichotomy Method

When crawling for POI data on the web, a rectangular geographic region, expressed by four coordinates, was taken as the input. In this region, the maximum tolerant crawling number of data was limited owing to the performance of the interface. In an actual scenario, the map structure is not regular, and the POI samples are not distributed uniformly. Hence, it is difficult to guarantee that the actual number of POI data in the selected rectangular geographic region exactly matches the maximum tolerant crawling number. If the configuration of a geographic region is too small or the POI samples are sparsely distributed, less than the maximum quantity of POI data is acquired for each crawling step. Consequently, the complexity of the data extraction algorithm will increase to obtain the fully sampled POI data. On the contrary, if the configuration of a geographic region is very large or the POI samples are densely distributed, the POI data may exceed the maximum tolerant crawling number. In this case, a few samples will not be successfully extracted, resulting in an incomplete data sample.

To solve the aforementioned problems, a POI data extraction optimization method based on dichotomy was proposed [14]. In this study, we constructed a two-dimensional coordinate system and defined the rectangular geographic region by four points, as shown in Eq. (1) and Fig. 2.

$$R = \{(x_1, y_1), (x_2, y_1), (x_2, y_2), (x_1, y_2)\} \tag{1}$$

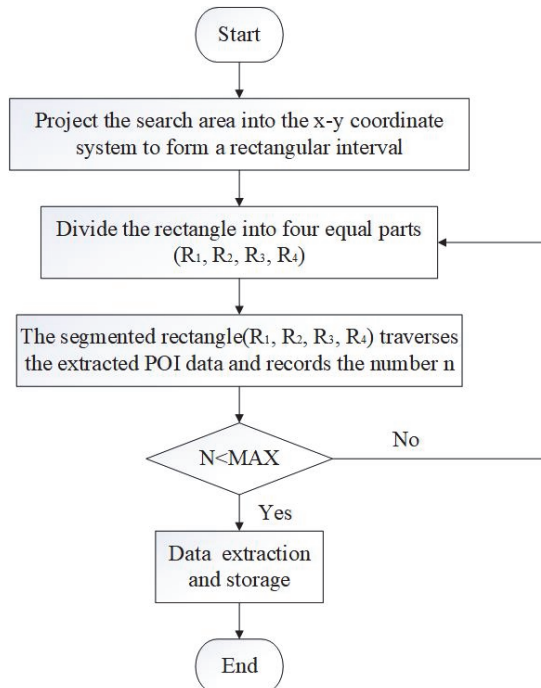


**Fig. 2.** Data extraction based on dichotomy method.

In Fig. 2, each node in the rectangle A-B-C-D denotes a full record of the POI data. In this study,  $N_{max}$  represented the maximum number of POI data for each step of the crawling operation. Initially, the rectangle was configured to completely cover the entire geographic area for analysis. In this case, the number of POI data significantly exceeded  $N_{max}$  and the rectangle was further divided to satisfy the data extraction demand. In the initial rectangle, the midpoints of the opposite sides were line-connected to divide the rectangle into quarters. The entire geographic area can be further expressed as a total of four sub-areas, as shown in Eq. (2).

$$R = \begin{cases} R_1 = \{(x_1, y_1), (\frac{|x_1 - x_2|}{2}, y_1), (\frac{|x_1 - x_2|}{2}, \frac{|y_1 - y_2|}{2}), (x_1, \frac{|y_1 - y_2|}{2})\} \\ R_2 = \{(\frac{|x_1 - x_2|}{2}, y_1), (x_2, y_1), (x_2, \frac{|y_1 - y_2|}{2}), (\frac{|x_1 - x_2|}{2}, \frac{|y_1 - y_2|}{2})\} \\ R_3 = \{(x_1, \frac{|y_1 - y_2|}{2}), (\frac{|x_1 - x_2|}{2}, \frac{|y_1 - y_2|}{2}), (\frac{|x_1 - x_2|}{2}, y_2), (x_1, y_2)\} \\ R_4 = \{(\frac{|x_1 - x_2|}{2}, \frac{|y_1 - y_2|}{2}), (x_2, \frac{|y_1 - y_2|}{2}), (x_2, y_2), (\frac{|x_1 - x_2|}{2}, y_2)\} \end{cases} \quad (2)$$

Further, each sub-area was traversed to check whether the actual number of POI nodes  $N_{R_i}$  matched the maximum tolerant crawling number  $N_{max}$ . If  $N_{R_i} \leq N_{max}$ , the data crawling operation was executed, and the POI data in the sub-area  $R_i$  were extracted. Otherwise, the rectangular sub-area was further divided into quarters based on the dichotomy method. When the entire geographic area to be analyzed was covered, the data extraction process was considered complete, and the data were stored for further analysis. The design flowchart of the algorithm is presented in Fig. 3.



**Fig. 3.** Flowchart of proposed dichotomy.

### 3. UFD Division Method based on Mean Shift Algorithm

Mean shift is a non-parametric and mode-seeking feature-space analysis method widely used in computer vision and image processing applications [15]. In the iterative procedure, each data point shifts to the average of its neighborhood based on the location distribution and attributive characteristics. A sample mean shift vector is described below.

Consider  $n$  data in a  $d$ -dimensional Euclidean space  $R^d$ ,  $x_i, i = 0, 1, \dots, n$ . The sample mean at  $x$  is calculated using Eq. (3).

$$M_h(x) = \frac{1}{k} \sum_{x_i \in S_h} (x_i - x) \tag{3}$$

Here  $S_h$  is a generalized  $d$ -dimensional sphere with radius  $h$ , as shown by Eq. (4):

$$S_h(x) \equiv \left\{ y : (y - x)^T (y - x) \leq h^2 \right\} \tag{4}$$

$k$  in Eq. (3) denotes the number of  $n$  data points located in the area of  $S_h$ .

Considering the weights of the distance between nodes and the linking of a node itself, the sample mean can be expanded as given in Eq. (5).

$$M_h(x) \equiv \frac{\sum_{i=1}^n G_H\left(\frac{x_i - x}{h}\right) w(x_i) (x_i - x)}{\sum_{i=1}^n G_H\left(\frac{x_i - x}{h}\right) w(x_i)} \tag{5}$$

In Eq. (5),  $w(x_i)$  is the weight function, and  $G_H$  is expressed by Eq. (6).

$$G_H\left(\frac{x_i - x}{h}\right) = |H|^{-1/2} G\left(H^{-1/2} \left(\frac{x_i - x}{h}\right)\right) \tag{6}$$

where  $G(x)$  is the kernel and  $H$  is a positive definite symmetric matrix, that is, the bandwidth matrix.

The mean shift algorithm uses two kernels:

The uniform kernel, as shown by Eq. (7).

$$F(x) = \begin{cases} 1 & \text{if } \|x\| < \lambda \\ 0 & \text{if } \|x\| \geq \lambda \end{cases} \tag{7}$$

and the truncated Gaussian kernel, as shown by Eq.(8).

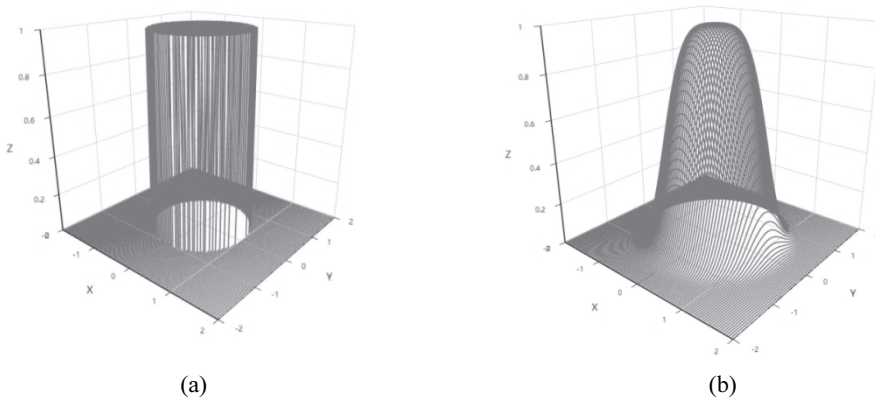
$$N(x) = e^{-\alpha \|x\|^2} \tag{8}$$

The value variations of the two kernels are presented in Fig. 4.

In UFD division, the geographical distance between two POIs presents a positive correlation with the possibility that they may be gathered together and classified into the same UFD. Hence, the shorter the distance from the clustering center, the larger the weight value that should be assigned. To achieve this, the truncated Gaussian kernel was selected for building the mean shift model in this study.

As presented in Table 1, urban POIs can be classified into 10 broad categories. However, no quantitative description is available for the attribute information, which has a decisive effect on UFD division. In [16], the authors proposed a method to extract hierarchical landmarks from urban POI data

according to their significant attributes. In that work, a significance measure model comprising three vectors, namely, public cognition degree, urban centrality degree, and characteristic attribute value, was constructed by analyzing the factors influencing the significance of POI objects from public cognition, spatial distribution, and individual characteristics. In this study, we refer to the aforementioned conclusions to build the UFD division model. The weight values  $w(x_i)$  for different types of POI were assigned as shown in Table 3.



**Fig. 4.** Unit uniform kernel function and unit Gaussian kernel function.

**Table 3.** Weight values for different POI data categories

Category	Weight value
Food service	0.5562
Shopping service	0.8146
Hotel	0.5562
Tourist attractions	0.6458
Residential building	0.0100
Medical service	0.5069
Leisure and entertainment	0.5010
Exercise and fitness	0.5010
Incorporated business	0.3057
Educational training	0.6706

Given kernel  $G(x)$  and weight  $w(x)$ , Eq. (1) can be further expressed as Eq. (9).

$$M_h(x) = \frac{\sum_{i=1}^n G\left(\frac{x_i - x}{h}\right)w(x_i)x_i}{\sum_{i=1}^n G\left(\frac{x_i - x}{h}\right)w(x_i)} - x \tag{9}$$

Let:

$$m_h(x) = \frac{\sum_{i=1}^n G\left(\frac{x_i - x}{h}\right)w(x_i)x_i}{\sum_{i=1}^n G\left(\frac{x_i - x}{h}\right)w(x_i)} \tag{10}$$

The UFD division method based on mean shift algorithm comprises the following steps:

**Initialization:** Each POI data node in the whole sample is defined as the initial point and minimum tolerance error  $\varepsilon$  is set as the convergence condition of the mean shift algorithm.

**Step 1:** Calculate  $m_h(x)$ .

**Step 2:** Assign  $m_h(x)$  to  $x$ .

**Step 3:** If  $\|m_h(x) - x\| < \varepsilon$ , finish the circulation. Otherwise, proceed to Step 1.

**Step 4:** Take each POI node as the initial point  $x$  and execute the preceding steps. Finally, the centers of the UFDs will be obtained.

## 4. Case Study

### 4.1 Performance of Dichotomy-based Data Extraction

To verify the performance of the proposed dichotomy method for POI data retrieval, the traditional equality division method, which divides the entire geographical area into several equal sub-areas, was selected as a contrast. The complexity of the two methods was calculated for two cases under the same sample data conditions. For Case 1, the two methods were used to extract all the data points in the case sample, and the required searches, which denote the algorithm complexity, were compared. For Case 2, the two methods were used to perform the same steps, and the final tally of extracted POI data points was compared. The results are presented in Table 4.

It is evident from Table 4 that the proposed dichotomy method can reduce the number of steps required for extracting all the data points in the given sample by 40% compared with the traditional equality division method. Moreover, with the same number of steps (228), the traditional equality division method obtained only 9,146 data points whereas the proposed dichotomy method extracted 12,615, an increase of 27%. It is concluded that the dichotomy method applied to POI data retrieval demonstrates better performance by reducing data extraction complexity, and simultaneously improves the data extraction efficiency.

**Table 4.** Analysis and comparison

Method	Number of steps (algorithm complexity)	Total number of POI
Dichotomy	228	12,615
Equality division		
Case 1	323	12,615
Case 2	228	9,146

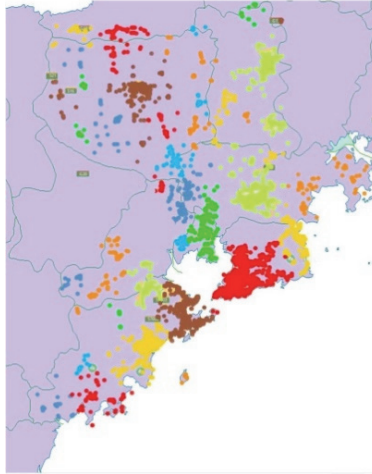
### 4.2 UFD Division Result Analysis

The case sample was crawled from the Internet, and the contained a total of 10 million nodes. All these points were printed on a map using JavaScript, as shown in Fig. 5. For an intuitive presentation of the points, those with different attributes were assigned different colors.

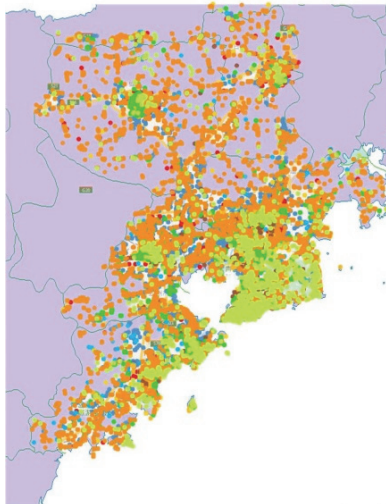
The results of the UFD division are presented in Fig. 6. All 10 million nodes were classified into 48 regions with different color expressions. It is obvious that our division results show evidence of



geographical agglomeration features. The division result also matches the UFD distribution in the official urban master plan (from 2011 to 2020), which is published by the city government and is quite authoritative (Fig. 7).



**Fig. 5.** Distribution of the POI nodes in the entire sample.



**Fig. 6.** UFD division result based on mean shift algorithm.

To further verify the performance of the proposed method, a quantitative similarity index was used to describe the rationality of the divided UFD. The similarity index was calculated using Eq. (11):

$$a = \frac{\sum_{i=1}^n x_i}{\sum_{i=1}^n X_i} \times 100\% \quad (11)$$

Here,  $n$  is the number of UFDs,  $X_i$  is the full score of the similarity of nodes in each UFD, and  $x_i$  is the actual similarity mark [17].

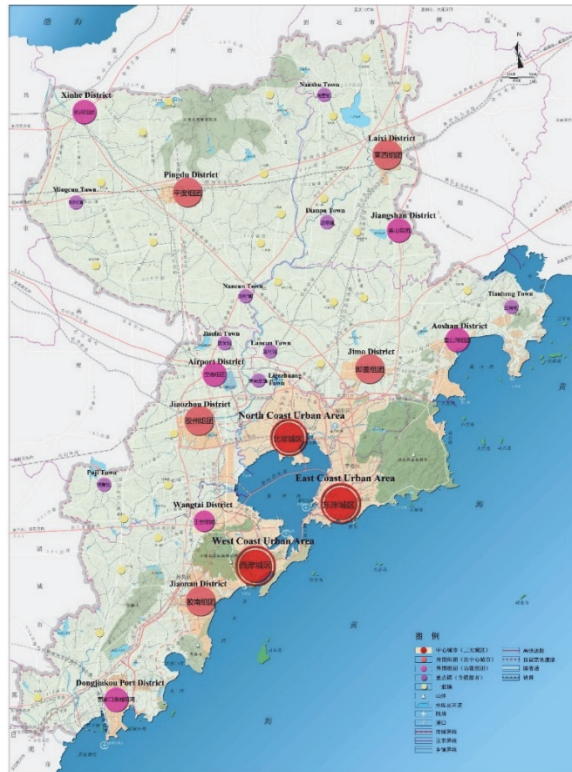


Fig. 7. UFD distribution in the official urban master planning.

We selected 23 UFDs from a total of 48 to verify the performance. The selected 23 UFDs contained a higher number of nodes and covered a relatively large geographical area, as shown in Table 5. In the table, the compliance score denotes the conformity of the nodes in the same UFD. The highest score is 3, which represents complete conformity. Scores 2, 1, and 0 denote comparative conformity, comparative inconformity, and complete inconformity, respectively.

Table 5. Conformity score of nodes in different UFDs

No.	UFD	Compliance score	No.	UFD	Compliance score
#01	East Coast Urban Area	3	#13	Jiangshan District	2
#02	North Coast Urban Area	3	#14	Xinhe District	2
#03	West Coast Urban Area	3	#15	Nanshu Town	3
#04	Laixi District	3	#16	Mingcun Town	2
#05	Pingdu District	3	#17	Dianpu Town	3
#06	Jimo District	3	#18	Nancun Town	3
#07	Jiaozhou District	2	#19	Tianheng Town	3
#08	Jiaonan District	3	#20	Jiaolai Town	2
#09	Aoshan District	3	#21	Lancun Town	1
#10	Airport District	3	#22	Ligezhuang Town	3
#11	Wangtai District	3	#23	Puji Town	3
#12	Dongjiakou Port District	2	-	-	-

We can observe from the table that the number of UFDs with complete conformity is 16, comparative conformity is 6, and comparative inconformity is 1. Different UFDs have different compliance scores. For UFDs with a high density of POI data samples, the division results coincide well with the actual scenarios. However, for certain UFDs, such as Lancun Town in Table 5, the compliance score is quite low. This is because only a few POI samples were generally extracted from the Internet for these UFDs. In such cases, the final clustering centers are most vulnerable to the distribution of these limited samples and can induce deviations. In total, the similarity index was 88.41%, referring to Eq. (11). In conclusion, the division results suitably match the actual scenario.

## 5. Conclusion

Currently, UFD relies on field and questionnaire surveys. This type of work is time-consuming and entails high human resource requirements. Moreover, the division of functional areas is greatly affected by subjective components, and with the acceleration in urbanization, the urban functional areas have become more complex and diverse, increasing the difficulty of field investigation. According to the research presented in this paper, the algorithm can automatically identify the functional area, and the clustering method can be proven. The demand data are simple to acquire and do not require heavy manpower and material resources. The results of the algorithm were relatable to the actual functional area. The distribution is quite similar, the operation is simple, and accuracy can be ensured. Our proposed method can provide information support for the decision-makers of the city to rationally formulate urban planning strategies and solve problems in the urbanization process.

## Acknowledgement

This paper is supported in partially by Shandong Provincial Natural Science Foundation (No. ZR2019QF017) and Basic Research Plan on Application of Qingdao Science and Technology (No. 19-6-2-3-cg).

## References

- [1] J. Wang, C. Li, Z. Xiong, and Z. Shan, "Survey of data-centric smart city," *Journal of Computer Research and Development*, vol. 51, no. 2, pp. 239-259, 2014.
- [2] H. W. Zhao, "The application of information surveying and mapping technology in national land survey," *Construction & Design For Project*, vol. 2019, no. 5, pp. 72-78, 2019.
- [3] Y. Li, S. Geng, X. Zhang, and H. Zhang, "Study of thermal comfort in underground construction based on field measurements and questionnaires in China," *Building and Environment*, vol. 116, pp. 45-54, 2017.
- [4] Q. Yin, S. Y. Zhu, and C. L. Gong, "Remote sensing analysis of the relationships between daytime ground bright temperature and land-use types of city: with shanghai as an example," *Journal of Infrared and Millimeter Waves*, vol. 28, no. 2, pp. 133-136, 2009.
- [5] Z. Chen and T. C. Hutchinson, "Probabilistic urban structural damage classification using bitemporal satellite images," *Earthquake Spectra*, vol. 26, no. 1, pp. 87-109, 2010.

- [6] Y. D. Eo, S. J. Moon, B. K. Lee, and B. W. Park, "Analysis of 3 D city image updating techniques using terrestrial photogrammetry," *International Journal of Digital Content Technology and its Applications*, vol. 6, no. 17, pp. 520-531, 2012.
- [7] D. J. Shane, M. A. Rufo, M. D. Berkemeier, and J. A. Alberts, "Autonomous urban reconnaissance ingress system (AURIS): providing a tactically relevant autonomous door-opening kit for unmanned ground vehicles," in *Proceedings of SPIE 8387: Unmanned Systems Technology XIV*. Bellingham, WA: International Society for Optics and Photonics, 2012.
- [8] Q. Yan, C. Li, C. Chen, and F. Luo, "Characteristics of activity space and community differentiation in Changchun: a study using mobile phone signaling data," *Human Geography*, vol. 33, no. 6, pp. 35-43, 2018.
- [9] X. Zhang, S. Du, and Q. Wang, "Hierarchical semantic cognition for urban functional zones with VHR satellite images and POI data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 132, pp. 170-184, 2017.
- [10] G. Pan, G. Qi, Z. Wu, D. Zhang, and S. Li, "Land-use classification using taxi GPS traces," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 1, pp. 113-123, 2013.
- [11] H. Zhang, R. Wang, B. Chen, Y. Hou, and D. Qu, "Dynamic identification of urban functional areas and visual analysis of time-varying patterns based on trajectory data and POIs," *Journal of Computer Aided Design & Computer Graphics*, vol. 30, no. 9, pp. 1728-1740, 2018.
- [12] Y. Wan and R. Wang, "Research on POI automatic classification assisted by comment information," *Journal of Geomatics*, vol. 43, no. 5, pp. 120-123, 2018.
- [13] X. Guan and Y. Zeng, "Research progress and trends of parallel processing, analysis, and mining of big spatiotemporal data," *Progress in Geography*, vol. 37, no. 10, pp. 1314-1327, 2018.
- [14] X. Chen, Y. Li, M. Lu, J. Lu, and W. Chen, "Implementation of K-means algorithm based on dichotomy," *Radio Communications Technology*, vol. 43, no. 6, pp. 37-40, 2017.
- [15] K. Fukunaga and L. Hostetler, "The estimation of the gradient of a density function, with applications in pattern recognition," *IEEE Transactions on Information Theory*, vol. 21, no. 1, pp. 32-40, 1975.
- [16] W. Zhao, Q. Li, and B. Li, "Extracting hierarchical landmarks from urban POI data," *Journal of Remote Sensing (Yaogan Xuebao)*, vol. 15, no. 5, pp. 973-988, 2011.
- [17] Y. Kang, Y. Wang, Z. Xia, J. Chi, M. Jiao, and Z. W. Wei, "Identification and classification of Wuhan urban districts based on POI," *Journal of Geomatics*, vol. 43, no. 1, pp. 81-85, 2018.



**Guo Hao** <https://orcid.org/0000-0001-9984-3291>

He received the bachelor's degree in Shandong Agricultural University, China, in 2018. Since September 2018, he is with the School of College of Geomatics from Shandong University of Science and Technology as a M.S. His current research interests include traffic big data



**Haiqing Liu** <https://orcid.org/0000-0003-4094-7541>

He received the bachelor's degree in automation from Central South University, China, in 2008, the Ph.D. in system engineering from Shandong University, China, in 2015. He is a lecturer in Shandong University of Science and Technology. From 2015 to 2017, he worked as a Post-Doctoral Researcher at the post-doctoral work station of Hisense Group, China. His current research interests include traffic engineering and control, cooperative vehicle infrastructure system and traffic intelligent perception.



**Shengli Wang** <https://orcid.org/0000-0002-8852-8621>

He received the Ph.D. degree in instrument science and technology from Southeast University, China, in 2013. He is an associate professor in Shandong University of Science and Technology. From 2015 to 2017, and is the chief engineer of Shandong Astro-compass Information Technology Co. Ltd. His current research interests include traffic engineering and control, cooperative vehicle infrastructure system and co-location.



**Yu Zhang** <https://orcid.org/0000-0002-5186-8585>

Since September 2016, he is with the School of college of Transportation from Shandong University of Science and Technology as a B.S. His current research interests include traffic engineering and control.