

## RESEARCH ARTICLE

# Urban Interior Boundaries Delimitation Respecting to Human Mobility

YU WEI<sup>1</sup> AND XI ZHAO<sup>2</sup><sup>1</sup>College of Computer Science and Technology (College of Data Science), Taiyuan University of Technology, Taiyuan 030024, China<sup>2</sup>School of Management, Key Laboratory of the Ministry of Education for Process Control and Efficiency Engineering, Xi'an Jiaotong University, Xi'an 710049, China

Corresponding author: Xi Zhao (1029198725@qq.com)

This work was supported in part by Shanxi Province Basic Research Program under Grant 202203021222103.

**ABSTRACT** Administrative divisions are regional divisions of the state for the purpose of hierarchical administration. With the development of urbanization process, a concern arising from the changing urban dynamics is that whether current administrative division accords with urban development? Following the concept of urban space formed by human activities, existing studies have sought proxies for human activities and then delineated urban boundaries formed by human activities from a data-driven perspective. However, their lack of exploration into the implicit human activity patterns behind the proxies results in weak interpretability and rationality of methods and results. This paper studies the delineation of urban interior boundaries respecting to human mobility patterns. Specifically, this paper first uses human mobility as a proxy to construct a network that represents spatial interaction in urban regions. Then we explore the region-based human crowd mobility patterns to reveal the mechanisms behind the formation of urban space through human activities. Finally, we employ the community detection technology to naturally delimit the urban interior boundaries formed by human mobility, and make a comparison with the official urban boundaries. Taking Xi'an in Shaanxi Province of China as an example, we conduct the delineation on the spatial interaction network based on the real mobility information of 24,770,715 mobile phone users in the whole city. We find that human mobility can establish a stable correlation between regions (or capture the objective correlations between regions), and the human crowd patterns are applicable for mining unusual urban regions from the perspective of anomaly detection, which are of great significance for understanding the urban spatial structure based on empirical evidence. In addition, in the final delineation result, some unexpected communities that are closely linked due to human activities appear from the results, and these findings help the urban planners re-examine the administrative division.

**INDEX TERMS** Administrative divisions, a data-driven perspective, urban interior boundaries, human mobility patterns, community detection.

## I. INTRODUCTION

The world is in the process of urbanization. The world's population is shifting from rural to urban areas. In 1960, 33.61% of the world's population live in urban areas, today the percentage is 55% (Figure 1) and this percentage is expected to increase to 68% by 2050.<sup>1</sup> The population migration leads to the expansion of urban areas. Except

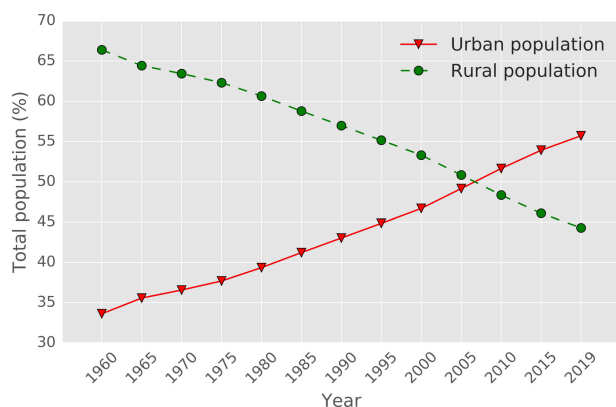
The associate editor coordinating the review of this manuscript and approving it for publication was Xianzhi Wang<sup>2</sup>.

<sup>1</sup><https://data.worldbank.org/topic/urban-development>

for these obvious changes, the urbanization has also led to economic and social changes [1], [2]. All these changes have impacts on the arrangement of spatial regions, which is known as the administrative division. The administrative division is committed to providing efficient management for the society, assisting in the overall development of the urban. Thus, facing the process of urbanization, it is necessary to re-examine the administrative division.

In reality, there have been cases in which the government tries to adjust administrative divisions, but the proposal has been opposed by citizens. This may be due to the subjective

nature of official adjustments [3], which do not align with the spatial cognition of urban residents. Spatial cognition is human understanding and perception of geographic space [4], [5]. For an individual, his spatial cognition is determined by his activities in the urban space [6], [7], and this spatial cognition dilutes the human impression of official boundaries. Respecting the boundaries of human activities is of great significance for urban management [8], [9]. Administrative divisions are the division of regions (small geographical spaces compared to urban) together. This is not a matter of individual circumstances, but needs to be in line with the spatial cognition of the relevant regional human crowd, which is influenced by the activity patterns of the regional human crowd. From a perspective consistent with human spatial cognition, urban space can be conceptualized as the environment created by human activity [10], [11]. Following this concept, this paper studies the delineation of urban interior boundaries respecting to human activities from a data-driven perspective.



**FIGURE 1.** Change in the proportion of urban and rural population worldwide from 1960 to 2019.

To achieve this, a complete and reasonable process is to first find a good proxy to characterize the human activity across spatial regions, and then explore the corresponding activity pattern of regional crowd, and finally try to divide the urban space according to the activity pattern. The region here is a small geographical space, which is the basic unit of urban space division. However, after finding the proxy for the human activity, existing studies [12], [13], [14], [15], [16] directly implements community detection methods on the spatially-embedded networks deriving from human activity across regions to delineate the urban boundaries. These studies lack an examination on the activity patterns of regional crowd, they do not explain what kind of regional crowd's activity pattern they follow and how their methods respect to the activity pattern, all of which lead no basis for the practice of their methods and no the interpretations for their results.

In fact, the proxies of existing studies are diverse, involving communication, trade, social contact and so on, and all

correspond to the complex and varied spatial interactions. Their lack of explanation may lead to a misunderstanding, that is, any activity can be used to establish a network to carry out community detection, and the results can be taken as the result of urban interior boundaries, which is obviously unreasonable. Administrative divisions should follow the activity pattern of regional human crowd, with the aim of grouping together regions that are relevant to the activities of regional human crowd [11]. However, human activities are diverse. An activity may vary greatly among individuals, so that the activities of a regional crowd may present an unpredictable situation. For example, the activities of a regional crowd may make the region establish a wide range of associations with other regions at random. At this situation, whether the corresponding human activity actually capture the objective spatial correlation between regions remains to be investigated, so the rationality of the results obtained by community detection technology is questionable. Even if community detection techniques are used to obtain partitioning results, the reliability and interpretability of the results are relatively poor.

Considering all of the above, this study employs the human mobility to characterize human activity across urban regions, explores human mobility patterns in a quest to reveal the mechanisms by which human activities form urban space, explores the feasibility of following human activities to delineate urban boundaries, and ultimately proposes a reasonable data-driven process for delineating urban boundaries.

This study conducts a case study to redraw the urban boundaries of Xi'an, China. At this time, the urban interior boundaries respecting to human activities should conform to the mobility patterns of regional crowd. Thus, we first explore the region-based human crowd mobility patterns to reveal the mechanism of spatial interaction. We characterize the region-based human crowd mobility patterns through three features: mobility entropy [17], [18], displacement [19], and human turnover, and find that the whereabouts of region-based human crowd present a relatively certain trend, which indicates that the human mobility establishes reliable correlations between regions. Then, we analyze the spatial interaction generated by human mobility from the perspective of network and anomaly detection respectively, and mine the key regions in the urban structure. These key regions contribute to a better understanding of human spatial interactions in urban. With these mobility patterns, we utilize the community detection technology in a logic way to redraw the interior boundaries of Xi'an. Compared with the official boundaries, several unexpected communities have emerged which really present closer internal relations. These findings capture the latest urban dynamics and help to objectively examine the urban interior boundaries.

The major contributions of this paper include the following:

- Our study sheds light on the mobility patterns of region-based human crowd, and concludes that the

human crowd moves within the urban in a relatively certain trend. These findings provide the evidence that human mobility can capture the objective correlation between urban regions and provide a reasonable basis for spatial division based on human mobility.

- In contrast to the network attribute-based approach commonly used in the literature, our study finds that anomalous mobility patterns of region-based human crowd can be used to explore key urban regions, which can help to reveal the spatial structure of urban.
- Based on the real human mobility data in a whole urban, we redraw the interior boundaries of Xi'an, China. Compared with the official urban boundaries, the results provide reasonable insights for urban planning.

We have organized this paper according to the following sections. Section II describes the work related to the determination of urban boundaries. In Section III, we introduce the process and main methods of this study. Specifically, in Section III-A, we first introduce the study area and research data, then we briefly describe the process of processing human mobility data into the spatial interaction network, which characterizes the regional interrelations. Then, in Section III-B, we explore the region-based human crowd mobility patterns, and in Section III-C, we describe the community detection method for automatically dividing urban regions. At last, in Section IV, we divide regions together in an automatic way to find the boundaries formed by human activities, and make a comparison with the official boundaries.

## II. BACKGROUND AND RELATED WORK

The administrative division is determined by government agencies to serve politics and administration. However, there may be several subjective operations in the process of determining the boundary [20], which may result in some improper divisions. For urban citizens, their perception of urban space depends on their activities in the urban [21], which weakens the impression for official boundaries. And for human activities in urban, modern research holds that socio-economic factors play an increasingly important role in driving human activities [22], [23]. Taking mobility in human activities as an example, its patterns may be changed by the socio-economic factors such as wage imbalance. Human activities are related to the social economy of urban, and can objectively reflect the dynamics of the urban to some extent. There have been studies exploiting human activities across the urban space to assess the effectiveness of urban growth boundaries [24]. Following these, Jiang and Miao [20] propose the 'natural city' which refers to the environment formed by human activities.

Under these circumstances, existing studies have sought to employ a good proxy to characterize the human activities across urban regions, and then resort to the network-based approach to detect objective urban boundaries. Chen et al. [12] utilize the activity trajectories of Nanjing citizens riding shared bicycles to construct a spatial interaction network,

and then implement a community detection algorithm on this network to delineate urban activity areas, and ultimately compare the boundaries of the areas with the boundaries of official administrative districts. Jin et al. [13] utilize the human travel data obtained from Baidu Huiyan Platform to conduct the spatial interaction analysis, and then directly apply the community detection algorithms to delineate the borders of activity spaces in the city. Shen and Batty [14] apply a multi-level modularity optimisation algorithm to detect community structures in the London Metropolitan Area, where the multi-level is designed for the commuting flows of different groups of people, so the method is essentially a hierarchical use of community detection algorithms directly on spatial interaction networks. Blondel et al. [25] employ mobile communications to examine the border in Belgium, and detect the underlying linguistic border. Sobolevsky et al. [15] employ human communication activities to characterize the spatial interaction within the country, such as France, and redraw the geographical maps according to human activities. Zhong et al. [16] use the smart card data of three years to respectively determine the boundaries of Singapore in these three years, and identify the evolution of urban structure brought about by development.

All these studies build spatial interactions through human activities first, and then directly apply the community detection technology to aggregate related regions. However, few studies have truly understood the patterns of these human activities, let alone how they follow the human activity patterns. Although human activities may follow the distance decay effect, that is, the intensity of interaction between two regions decreases with the increase of geographical distance, the effect does not explicitly express the correlation between regions, and only this effect is insufficient to explain how human activities shape the boundaries of 'natural city'. The patterns of human activities are not clear, so the rationality of using community detection technology directly is questionable, and the interpretability of corresponding results is poor. Therefore, this study first explore the mechanism of spatial interaction produced by human mobility, making sure that this spatial interaction do establish reliable correlations between regions, and then we delineate the urban boundaries respecting to human activities.

## III. MATERIALS AND METHODS

The framework of this study mainly contains three parts (Figure 2). The first part is the data preparation stage. We map individual raw cellular data into urban spatial activity trajectories, and then aggregate crowd trajectories to form spatial interaction network. The second part is the analysis of region-based human crowd mobility patterns. At this stage, three mobility-related features mobility entropy, displacement, and human turnover are proposed. We analyze the basic human crowd mobility patterns through univariate analysis and correlation analysis, and then conduct anomaly pattern analysis to explore the abnormal regions. The third part is urban interior boundaries delimitation. Community

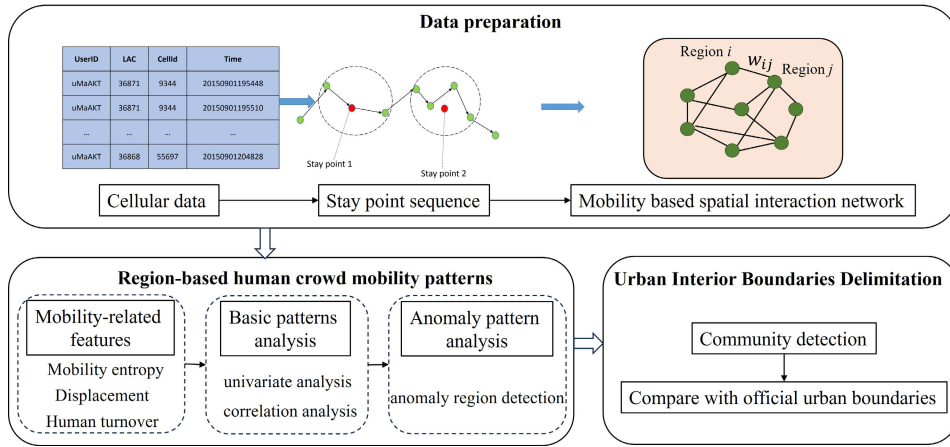


FIGURE 2. Framework of the methodology.

detection algorithm is applied to automatically delineate the internal boundaries of urban, and the obtained boundaries are used for comparison with the actual official boundaries.

## A. STUDY AREA AND DATA PROCESSING

### 1) XI'AN DISTRICTS

As a national central city in western China, Xi'an is an important scientific research, education and industrial base in China. As of 2023, Xi'an covers an area of 10752 square kilometers, and comprises 11 districts and 2 counties.

In history, the administrative division of Xi'an has been adjusted many times. Administrative division aims to serve the development of urban, regional economy and society. In recent years, Xi'an has experienced rapid development. The urban population has expanded from 8.156 million in 2015 to 9.8687 million in 2018; the built-up area of Xi'an has grown from 548.60 square kilometers in 2015 to 724 square kilometers in 2018; and the urban GDP has increased from 58.12 billion in 2015 to 83.4986 billion in 2018.<sup>2</sup> In order to satisfy the development needs of national central cities and better undertake the task of leading regional economic development, it is necessary to examine the existing administrative divisions of Xi'an. The study area covers all the districts of Xi'an, as shown in Figure 3.

### 2) CELLULAR DATA

Our research data, coming from the 3G cellular network of Xi'an, records the mobility information of 24,770,715 mobile phone users in Xi'an from October 1, 2015, to October 31, 2015. A mobile phone in normal operation will be associated with the nearest base station to record its current location, leaving the original location data.

We process each user's above daily original data into stay point sequence which represents this user's space activity trajectory of that day referring to studies [26]. In our study, a stay point refers to a geographical region within 500 meters

in which a user stays in this region for more than 30 minutes. Stay points indicate the meaningful activity regions for users.

Then, a user  $i$ 's space activity trajectory in one day is represented as a stay point sequence  $SP_i = \{sp_{i1}, sp_{i2}, \dots, sp_{in}\}$ , where each point  $sp_{ij} \in SP_i$  contains the longitude, latitude and arrival and departure time.

### 3) MOBILITY BASED SPATIAL INTERACTION NETWORK

Based on the stay point sequence in section III-A2, we obtain the spatial interaction information generated by human mobility. Specifically, taking a user's stay point sequence as an example, the spatial activities of this user result in the spatial interactions between corresponding regions of the adjacent stay points. Considering the spatial activities of all users, we can obtain the representation of spatial interaction among the regions within the urban caused by human mobility. We define the interaction network as an undirected weighted graph  $G = (\mathbb{N}, \mathbb{E}, \mathbb{A})$ , where  $\mathbb{N}$  contains all the regions in the study area,  $\mathbb{E} = \{(i, j) : i, j \in \mathbb{N}\}$  contains all edges which directly connect the regions  $i$  and  $j$ , and  $\mathbb{A} = \{A_{ij} : i, j \in \mathbb{N}\}$  are the travel volume between the regions  $i$  and  $j$ .

According to the data processing in section III-A2, adjacent stay points may relate to the same geographic region. Our focus is on the interaction between regions. Therefore, considering the large number of all users' stay points, we divide the urban into grids, and take grids as the basic region to study the spatial interaction.

To determine the grid size, with the above user trajectory data, we calculate each user's radius of gyration which is a metric to distinguish users' mobility patterns, and chose the distinct geographic distance which separates total users equally into two main groups as the grid size [27], [28]. Figure 4 depicts the cumulative distribution function of the radius of gyration, in which users with the radius of gyration less than 1094 meters account for 50.2% of the total. Finally, we set each grid as a square with 1000 meters width. Mapping

<sup>2</sup><http://tjj.xa.gov.cn/tjnj/2019/zk/indexch.htm>



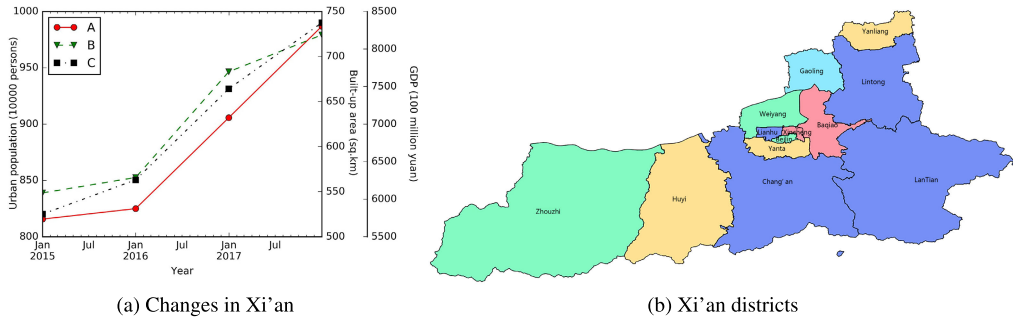


FIGURE 3. The study area.

the above user stay points into grids, we calculate the travel volume between any two grids.

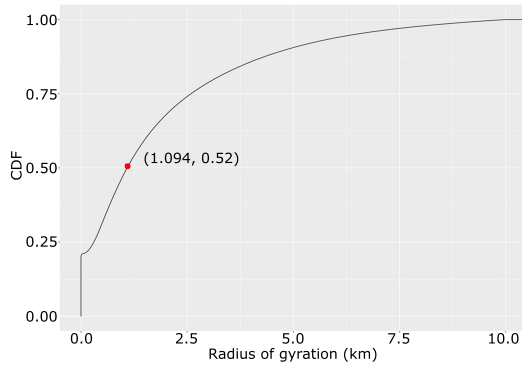


FIGURE 4. The CDF of the radius of gyration.

Then, we obtain the interaction network generated by human mobility. Descriptive statistics for the basic topological properties of the interaction network are provided in Table 1.

Based on the above introduction, we formulate the problem as follows: given the mobility based spatial interaction network  $G$ , divide it to form the urban interior boundaries, respecting to human mobility patterns.

TABLE 1. The basic topological properties of the interaction network.

Attribute	Value
Number of nodes	3325
Number of edges	1276913
Average degree	768.068
Average trip volume by weighted edges	72323.67
Average shortest path length by edges	1.783

## B. REGION-BASED HUMAN CROWD MOBILITY PATTERNS

Although we have obtained the urban spatial interaction network, the regional crowd mobility pattern that affects the spatial division is still unclear. If the mobility of region-based human crowd presents an irregular situation, it is impossible to delineate the reliable boundaries respecting to human

activities. Therefore, it is imperative to explore the human crowd mobility patterns to better understand the spatial interaction caused by human mobility.

### 1) BASIC PATTERNS ANALYSIS

In this section, we first explore the entropy and displacement of region-based human crowd. Then coupling with the human turnover which is measured by the associated travel volumes, we comprehensively understand the urban spatial interaction.

In the study of individual mobility patterns, [17] measure the entropy of each individual's trajectory, and find the potential predictability of human mobility. Here, the entropy is a measure of uncertainty [29], [30]. In our study, the spatial interaction is related to the human mobility at collective level. Therefore, we take the region-based human crowd as the basic unit, and treat the spatial interactions between regions as the representation of human crowd mobility, and then employ the entropy to measure the certainty of human crowd mobility. Taking region  $i$  as an example, regions interacting with region  $i$  are considered as the locations visited by the human crowd in region  $i$ , and the corresponding travel volumes measure the interaction frequencies with region  $i$ .

We employ two types of entropy to measure the region-based human crowd mobility pattern. The first random entropy is calculated as:

$$H1 = \log_2 n_i, \quad (1)$$

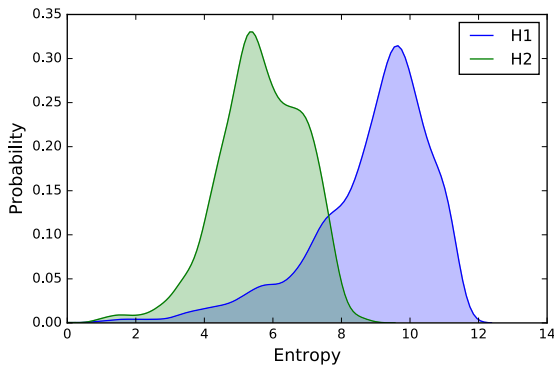
where  $n_i$  is the number of regions interacting with region  $i$  through human mobility, indicating the probability of the human crowd in region  $i$  randomly interacting with other regions. The second region-based human crowd mobility entropy is calculated as:

$$H2 = - \sum_{j=1}^{n_i} p_i(j) \log_2 p_i(j), \quad (2)$$

where  $p_i(j)$  is the probability of the human crowd in region  $i$  interacting with region  $j$ , characterizing the uncertainty of the human crowd interaction in region  $i$ .

To characterize the uncertainty of human interaction across the whole regions, we calculate  $H1$  and  $H2$  for human crowd in each region, and the distributions  $P(H1)$  and  $P(H2)$  are

shown in Figure 5. There is a obvious left shift of  $P(H2)$  compared with  $P(H1)$ , which indicates that the mobility of region-based human crowd is far from random.  $P(H1)$  peaks when  $H1$  is approximately equal to 9.5, which indicates that the human crowd in one region may on average interact with  $2^{H1} \approx 724.08$  regions, in the case of random movement. In contrast,  $P(H2)$  peaks when  $H2$  is approximately equal to 5.2, which indicates that the real uncertainty of the human crowd mobility in a typical region is not 724.08 but  $2^{5.2} = 36.76$ . These results demonstrate that region-based human crowd tends to interact with several certain regions, that is to say, there is a relatively directional trend in the mobility of human crowd.



**FIGURE 5.** The probability distribution of two mobility entropy. The  $H1$  is the random entropy, and the  $H2$  is the region-based human crowd mobility entropy. The obvious left shift of  $P(H2)$  compared with  $P(H1)$  indicates that the mobility of region-based human crowd is far from random.

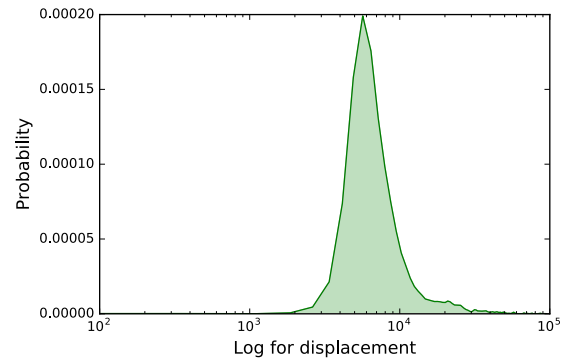
Except for the entropy, we explore the displacement of region-based human crowd, as it measures how far the human crowd typically moves. For the human crowd in region  $i$ , the displacement is calculated as:

$$d_i = \frac{1}{W_i} \sum_{j=1}^{n_i} w_{ij} \times d_{ij}, \quad (3)$$

where  $n_i$  is the number of regions interacting with region  $i$  through human mobility,  $w_{ij}$  is the travel volume between the regions  $i$  and  $j$ ,  $W_i = \sum_{j=1}^{n_i} w_{ij}$ , and  $d_{ij}$  is the distance between region  $i$  and region  $j$ .

Figure 6 shows the distribution of displacement. The distribution peaks when the displacement is approximately equal to 3162.68 meters. Thus, the movement range of human crowd in a typical region is within 3162.68 meters, which seems to indicate that there is an implicit boundary of crowd activity.

The above statistical results of the univariate analysis is a mixture of all regions with different situations. Taking into consideration the region human turnover which is measured by the associated travel volumes, we characterize each region with three features: entropy, displacement and region human turnover. To better understand the spatial interaction within



**FIGURE 6.** The probability distribution of displacement.

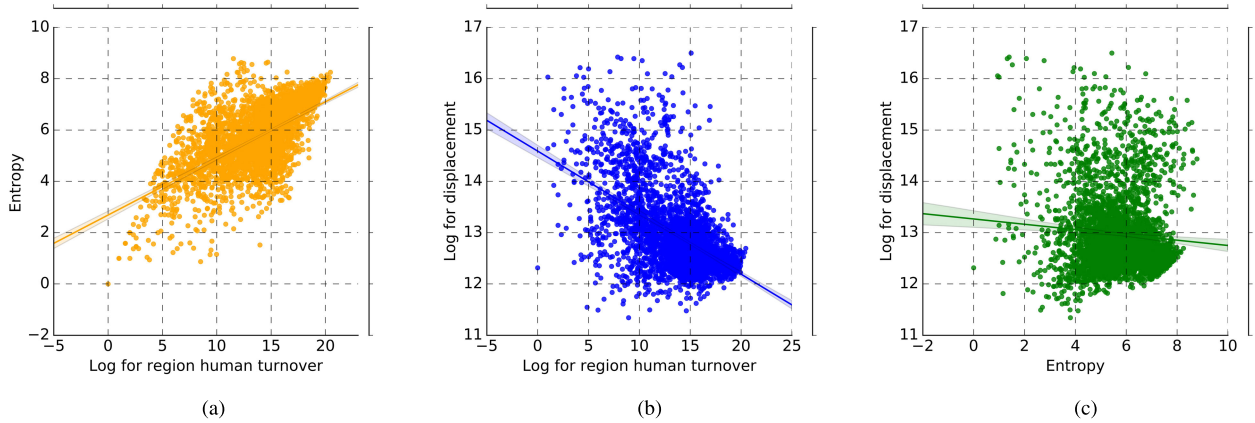
the urban, we conduct correlation analysis on these features to reveal more general conclusions next.

Figure 7a depicts a scatter plot of the region-based human crowd mobility entropy and the log amount of region human turnover. The plot overlays a linear trend, which presents a positive correlation between the entropy and region human turnover. This situation indicates that with the enhancement of region human turnover, the uncertainty of the human crowd's whereabouts also increases correspondingly. There are differences between individual mobility, and it is the diversity of a large number of individuals that leads to increased uncertainty at the collective level. While Figure 7b presents a negative correlation between the log of displacement and the log amount of region human turnover, which indicates that human crowd in the regions with high human turnover are more likely to interact with nearby spaces in close proximity. Combining these two results, we find that although the uncertainty of human crowd mobility is high in the regions with large human turnover, these crowds interact more evenly with the nearby regions around them; and in the regions with low mobility uncertainty, these crowds are more inclined to interact with some specific regions far away. These conclusions are further confirmed by Figure 7c.

In conclusion, although individual mobility presents diversity, at the collective level, region-based human crowd mobility exhibits a degree of certainty that is far from random, and the movement range seems to be limited by an implicit boundary. The region-based human crowd either interacts with neighboring regions, exhibiting localized activity, or are more inclined to interact with particular regions. All of these indicate that some regions are indeed more closely linked by human activities, and that this relationship is reliable, which reflects the objective interrelations between urban regions.

## 2) ANOMALY PATTERN ANALYSIS

The above results represent the major patterns of urban spatial interaction. Although these patterns are generic to most regions, we still notice that minority regions present anomalous patterns, which are presented as the outliers in Figure 7, such as the region with high region human turnover



**FIGURE 7.** Relationship between mobility entropy and displacement of region-based human crowd, and region human turnover. (a) The plot of mobility entropy and the region human turnover. (b) The plot of displacement and region human turnover. (c) The plot of displacement and entropy.

and small entropy in Figure 7a. Anomalies present the special patterns which are different from normal instances [31], [32]. In an urban environment, anomalies are sometimes critical, as they may undertake special functions or indicate the particular events in the urban [33], [34], [35]. The anomalies may play an important role in understanding urban interaction. Existing studies usually detect the key elements of urban structure, such as urban centers, from a network perspective to understand the urban interaction [16], [36], [37]. Here, from an anomaly detection perspective, we discriminate the outliers, and compare them with the key elements of urban structure mentioned in existing studies to better understand their role in urban spatial interaction.

Specifically, we use three features, entropy, displacement, and human turnover, to characterize the crowd mobility patterns in each region, and apply the Isolation Forest method to separate the anomalous instances from the rest of the instances [38]. This method explicitly isolates the anomalies based on the features, and ensure the feature-values of the separated anomalies are very different from the normal instance. Considering the complexity of human mobility, the spatial interactions in regions with low human turnover have a certain contingency, thus this study focuses on the anomalous regions with high region human turnover. Here, we present several anomalous regions and their corresponding features in Table 2. In these regions, the region human turnovers are relatively high, while the human crowd mobility entropy are small, and the displacements are large. This indicates that the large-scale human crowd in these regions tend to interact more directionally with several remote regions, which is contrary to most other regions.

**TABLE 2.** The features of several anomalous regions.

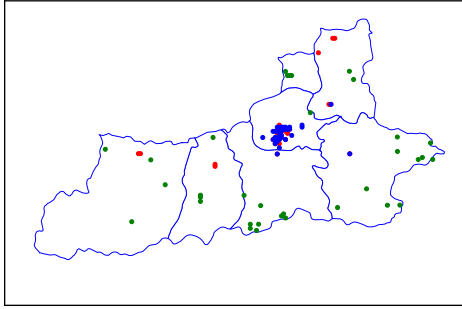
Region	Related travel volume(log)	Mobility entropy	Displacement(log)
1	13.97	1.52	16.27
2	13.32	2.77	16.35
3	10.98	1.16	11.81
4	9.46	1.32	16.39

Different from the above anomalies, existing studies mine the key elements of urban structure such as hubs from a network perspective to better understand urban interactions. These key elements are more central than others, and are often associated with the large degree property in the urban related network [39]. Next, we explore the relationship between the abnormal regions mined from our anomaly detection perspective and the key regions mined from the network perspective. Referring to the study [36], we employ two centrality indices betweenness centrality and PageRank to determine the hubs and centers in the urban, and compare them to our outliers.

According to the urban spatial interaction network, we calculate the two central indicators of each region, and select the first 1% of the two indicators to map to the geographical space. As shown in Figure 8, the hubs and centers are usually concentrated within the urban, and sometimes they are the same or very close to each other, while our outliers are located more on the outskirts of the urban. We further map the spatial interactions of some anomalous regions into geographical space as shown in Figure 9. We find that the large numbers of human in these anomalous regions are more likely to interact with the distant regions within the urban, such as the urban center regions. These outliers are more like the gateway for external people to enter the urban. For example, in Figure 9a, there is a bus station near the abnormal region, and in Figure 9b, there is a highway station entrance near the abnormal region. They may be the key elements for human commuting between the urban and the outside, and are important for understanding urban interactions. Together with the results mined from the network perspective, these elements mined from the human crowd mobility patterns help us better understand the urban structure.

### C. COMMUNITY DETECTION

According to the conclusions in Section III-B, human crowd in a typical region is more inclined to interact with several



**FIGURE 8.** The hubs(blue), centers(red) and anomalous regions(green). The hubs and centers are mined through the two centrality indices betweenness centrality and PageRank from a network perspective, and the anomalous regions are distinguished according to the region-based human crowd mobility patterns.

specific regions, and tends to move within a certain range. The interactions between regions are mutual, thus human crowds interact more purposefully within a certain range, which presents that the intra-interaction of the collection of regions within this range are much higher than the inter-interaction between this collection and other geographic regions. As a result, several regions are more closely linked through human activity, and it seems that these regions appear to be naturally isolated from other regions, just like there is a geospatial boundary formed by human activities. Therefore, we can delineate the boundaries formed by human activities through dividing closely related regions together. We hope to achieve this in an automatic way, and make a comparison with the official boundaries.

In our study, the mobility based spatial interaction network is the representation of human interacting across the urban regions, in which the nodes represent the regions and the weights represents the interaction intensities between regions. According to the patterns of human activity described above, dividing closely related regions together is consistent with the paradigm of community detection. Thus, based on the derived spatial interaction network, we utilize the community detection algorithm to determine the boundaries formed by human activities.

There are various community detection algorithms, and this research adopts the Louvain method [40] because of the interpretability of its algorithmic implementation process for community merging on spatial interaction network. The core of the Louvain method is modularity optimization [41], [42]. The modularity is defined as the number of edges falling within groups minus the expected number in an equivalent network with edges placed at random [41], and it is a commonly used metric to measure the strength of network community structure. For a particular division of the network, the modularity is expressed as

$$Q = \frac{1}{2m} \sum_{C \in \mathbb{P}} \sum_{i,j \in C} \left( A_{ij} - \frac{k_i k_j}{2m} \right) \quad (4)$$

where  $\mathbb{P}$  is the community set of a particular network division, node  $i$  and node  $j$  belong to the same community

$C$ ,  $A_{ij}$  is the number of edges between node  $i$  and node  $j$ ,  $m = \sum_{i,j} A_{ij}/2$ ,  $k_i$  and  $k_j$  are the degrees of node  $i$  and node  $j$  in the network, respectively. The modularity with a higher value indicates a stronger community structure of the network division, that is, the better the division quality is. The implementation details of the Louvain method are shown in Algorithm 1.

---

**Algorithm 1** Implementation of the Louvain Method in Spatial Interaction Network

---

**Input:**

Spatial interaction network  $G$ , travel volume matrix  $\mathbb{A}$ ;

**Output:**

Communities;

```

1: for each node  $i$  in  $G$  do
2:   assign  $i$  to its own community;
3: end for
4: while there is an improvement in modularity do
5:   for each node  $i$  in  $G$  do
6:     compute gain in modularity in Equation (4) for
       moving  $i$  to each neighbor community;
7:     if moving  $i$  to a neighbor community increases
       modularity then
8:       move  $i$  to the neighbor community with highest
       gain;
9:     end if
10:  end for
11:  create a new graph  $G'$  where nodes are communities
       from  $G$ ;
12:  for each community  $C$  in  $G'$  do
13:    set weight of edge  $(C, C')$  as sum of weights of
       edges between  $C$  and  $C'$ ;
14:  end for
15:   $G = G'$ .
16: end while
17: return Communities;

```

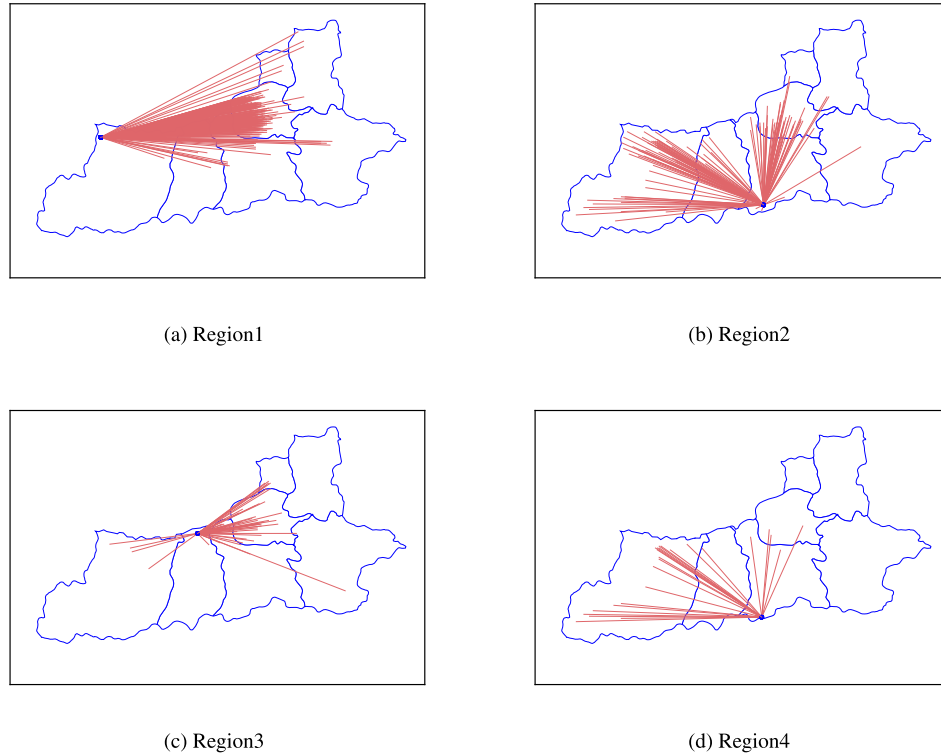
---

## IV. RESULTS AND DISCUSSION

### A. URBAN INTERIOR BOUNDARIES DELIMITATION

Urban boundaries delimitation that respects to human mobility is intended to divide regions together that are closely related by human mobility, so that regions within the boundary have less interaction with outside regions. According to the analysis in Section III-C, community detection technology is applied to spatial interaction network to delineate urban boundaries, and naturally the modularity is used to measure the quality of the delineation results, since modularity characterizes exactly the tightness of the community structure formed by the boundaries. In addition, some common sense of administrative divisions can be used to judge the quality of results, such as the community should not be too small or scattered.





**FIGURE 9.** The spatial interactions of four selected anomalous regions.

### 1) DETAILS OF DATA PROCESSING

Here, we illustrate the process of extracting human mobility information from cellular data.

For each user, his time-ordered mobility behavior record data can be represented as the sequence  $Rec = \{re_1, re_2, \dots, re_n\}$ , where each record  $re_i$  is of the form  $\langle AnonymizedID, Time, Lacci, longitude, latitude \rangle$ . The sequence is processed as follows:

- Exception record handling. Delete records that drift very far in a short period of time.
- Stay point extraction. Converge the adjacent records with spatial distances spanning within 500 meters and time spans of more than 30 minutes, and take the average of their latitude and longitude to obtain the stay point.

At this point, we obtain the stay point sequence for each user representing his semantic activity information. Then, the mobility based spatial interaction network is constructed according to Section III-A3.

### 2) BASELINES AND PERFORMANCE EVALUATION

Four representative community detection methods are used for urban boundaries delimitation.

- Louvain [40]. This method directly aims to optimize modularity by attempting different merge pairings of neighboring nodes in the network to achieve the maximum modularity possible.
- Infomap [43]. This method is based on the principle of minimum entropy. It first generates a sequence

through random walks, and then optimizing community partitioning using hierarchical encoding and entropy minimization.

- Walktrap [44]. The core idea of this method is to utilize the behavioral properties of random walks on the graph, that is, the walking process tends to fall into densely connected subgraphs (i.e. communities). In this way, a distance metric between vertices and between communities can be computed, which in turn leads to a hierarchical community structure through a clustering algorithm.
- Spectral clustering. This method uses the eigenvectors of the Laplacian matrix of the graph to segment the graph, and the community partitioning depends on the clustering results of the eigenvectors.

The modularity in Equation (4) is used as a measure of the closeness of associations within the boundaries, and some common sense that administrative regions should not be too small and scattered is used to judge the quality of the delineation results.

### 3) RESULT ANALYSIS

Among the four methods mentioned above, except for the spectral clustering that requires specifying the number of communities, the other methods automatically partition to obtain community results. The first three methods are conducted using the i-graph package on the R platform, and the spectral clustering is implemented through Python. The

**TABLE 3. The division results of three community detection methods.**

Method	Louvain	Infomap	Walktrap	Spectral clustering
Average number of communities	10	78	233	20
Average modularity	0.51	0.49	0.45	0.46

average number of communities and modularity of these methods after 100 times of execution are shown in Table 3, and here we choose to show the best results of the spectral clustering, that is, when the number of communities is set to 20.

In terms of the modularity, the Louvain achieves the highest score, followed by the Infomap and the Spectral clustering, and finally the Walktrap. The Louvain directly optimizes the modularity for community partitioning, whereas the other methods are based on an indirect criterion for partitioning, so the results are obvious. Remember that our original intention was to divide regions that are closely connected due to human activities together, making communities that are closely connected internally and their members less connected externally, so directly optimizing the degree of modularity is a good choice. The Infomap and Walktrap expect to obtain the structure of subgraphs through random walks, and then further obtain community partitioning by minimizing information entropy and clustering based on node distance metrics, respectively. Although capturing subgraph structures through random walks has some interpretability, they do not directly focus on the degree of connectivity within the community, and their results on modularity are generally taken for granted. As for spectral clustering, it cannot automatically obtain the number of regional divisions, and in addition, it obtains results through eigenvectors clustering, so its interpretability is the worst. Meanwhile it does not have a significant advantage in terms of the modularity.

As for the general common sense of administrative divisions, we explore the number of community divisions and the size of community composition. The Louvain divides the city of Xi'an into 10 communities, the Infomap results in 78 communities, and the Walktrap results in 785 communities. Compared with the 13 official administrative districts in Xi'an, the 10 communities in the Louvain are relatively close, while the other two methods Infomap and Walktrap generate too many communities leading to fragmentation, which is inconsistent with common sense. As for the size of community composition, note that the elements in the community, which is our basic research unit, are 1000 meters by 1000 meters sized urban regions. The number of elements in the 10 communities of Louvain are 569, 473, 451, 424, 316, 268, 262, 217, 193, 152 respectively. However, the other three methods generate a large number of single-element communities, which is inconsistent with reality. Among them, the spectral clustering is particularly excessive, which generates one larger community, while the other communities are small.

Considering the above results, we adopt the Louvain method. Specifically, this method employs a repeated two-stage process to divide the nodes into groups. At the first stage, each node is assigned to a community. At the second stage, for each node in the network, this method traverses all neighbor nodes of this node, and measures each gain of modularity brought by removing this node from its own community to the community where its one neighbor node belongs to. Then this node is assigned to the community which achieves the maximum modularity. This process traverses all nodes and is repeated until no higher modularity can be achieved. The whole process is interpretable and consistent with the original intent of our community detection. As for the computational complexity of the Louvain method, in the first stage, each node needs to compare with all adjacent nodes and calculate the modularity gain, thus the time complexity of this stage is  $\Theta(N^2)$ , where  $N$  is the number of nodes. In the second stage, node merging will reduce the number of nodes in each iteration, so the subsequent computational is gradually reduced. Due to the low complexity of merging operations and reconstructing the network, it usually does not significantly increase the overall computational complexity.

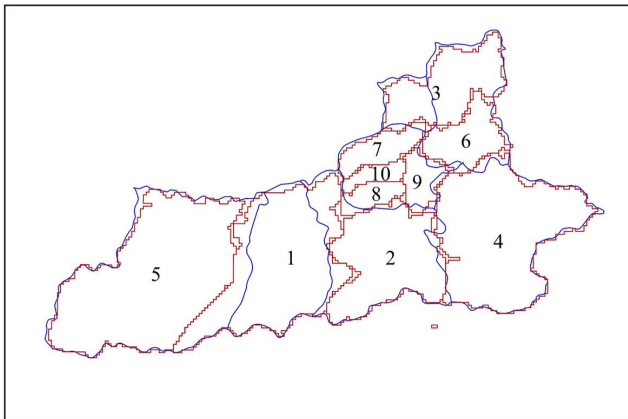
To further evaluate the quality of the network division, we calculate the internal interaction ratio for each community, i.e. the percentage of the travel volumes human move within the same community. As shown in Table 4, the internal interaction in each community achieves a high percentage, with 88.52% at the most and 50.74% at the least, which fulfills our purpose of dividing closely interacting regions together.

**TABLE 4. The internal interaction ratio of each community.**

Community number	Internal interaction ratio	Community number	Internal interaction ratio
1	82.13%	6	81.31%
2	65.33%	7	52.01%
3	88.52%	8	54.54%
4	87.38%	9	51.03%
5	87.40%	10	50.74%

## B. COMPARE WITH OFFICIAL URBAN BOUNDARIES

In this section, we make a comparison between the urban boundaries generated by human activities and the official urban boundaries. The boundaries of communities generated by human activities are shown in red in Figure 10, in which the blue boundaries are the official urban boundaries. Our main findings are as follows:



**FIGURE 10.** The 10 communities generated by human activities.

The urban space of Xi'an is divided into 10 regional sets, and all communities are made up of geographically adjacent regions. The community detection algorithm neither specifies the number of communities nor adds the restriction on the contiguity of regions. Only driven by optimizing the modularity, a measure of community quality, the method automatically obtains the community structure in which the members in the same community are closely connected, while the links between the communities are relatively less. According to the above human crowd mobility patterns, the region-based human crowds interact within a certain range, and some regions within the range may be naturally isolated from other regions due to the frequent human interactions. Therefore, it is reasonable for adjacent regions to be divided together, as shown by the community detection results. In real-world applications, it is common sense that geographically neighboring regions form a unified administrative region, thus the direct application of community detection algorithms without distance constraints may lead to irrational divisions. However, the exploration of human mobility patterns in this paper indicates that neighboring regions are more likely to form strong connections in the graph. Then in the specific implementation process of community detection algorithms, strongly correlated neighboring nodes in the graph are partitioned together, implicitly imposing a distance constraint.

Compared with the official 13 districts, the community 1,2,4 and 5 generally follow the official districts, while other communities differ greatly from the official districts. Among the four communities which are the most similar to the official districts, the boundary of community 4 is basically consistent with that of the official LanTian County, and the boundary of community 1 extends to the left and right, occupying part of the official administrative regions corresponding to community 2 and community 5, especially the northeast corner of community 1. Although the boundaries of community 1, 2, and 5 are somewhat offset from their official counterparts, the average internal interaction ratio of these three communities is 78.29%,

which is close to the official 79.56%, based on Table 4 and Table 5, and the internal interaction ratio of community 1 is even higher than that of official Huiyi district. Human space activities will naturally link some regions more closely to form so-called communities, and the connections between communities are relatively less, just like there are barriers. These barriers determine the boundaries of human activities, and these boundaries may be a more reasonable division of urban space. The consistency of some community boundaries with official boundaries suggests that human mobility in these regions follows official boundaries.

Then, we analyze the results that are clearly inconsistent with official boundaries. In the northeast corner of Xi'an, the three official districts Yanliang district, Lintong district and Gaoling district are merged into two communities: community 3 and 6. Specifically, the south of Lintong district alone forms community 6, while the rest of the three districts are merged to form community 3. According to Table 4 and Table 5, the average internal interaction ratio of community 3 and 6 is 84.92%, which is higher than 81.59% of the three corresponding official districts. Among them, the internal interaction ratio of community 3 is as high as 88.52%, which is not decreased due to regional merger, but higher than that of a single administrative district, especially the 74.29% in Gaoling district. In the case of these three official districts, the more interconnected regions are divided together by merging and regrouping. Our division organizes regions in a way that is more consistent with human activities, which is likely to reduce the urban management problems caused by a large number of cross-district activities.

For the remaining six official districts, our method divides them into four communities: community 7,8,9 and 10. Although the internal interaction ratios of these four communities are lower than that of community 1-6, their average internal interaction ratio is 52.08%, which is higher than 47.89% of the six corresponding official districts, according to Table 4 and Table 5. We further analyze the human interaction in these four communities. Here, we calculate the percentage of the interaction between each community and the other three communities, and the percentage of the interaction between each community and community 1-6. As shown in Table 6, except for the interaction with itself, the interaction between each community and the remaining three communities accounts for the vast majority. Taking the community 7 as an example, 52.01% of the interactions occur within itself, followed by 32.95% of interactions with the other three communities, and only 15.03% of interactions with community 1-6. The average internal interaction ratio is higher than that of the official districts, which means that our division is more consistent with human activities. As for the relatively low internal interaction ratio compared to community 1-6, it is just because of the more interaction between these four communities. As the core regions of Xi'an, these regions are highly integrated, which results in a large interaction between these regions and makes it difficult to divide the communities.

**TABLE 5.** The internal interaction ratio of each official region.

Official district	Internal interaction ratio	Official district	Internal interaction ratio
Huyi District	81.95%	Weiyang District	64.38%
Xincheng District	38.05%	Chang'an District	66.33%
Lianhu District	39.30%	Gaoling District	74.29%
Baqiao District	56.93%	Yanliang District	87.80%
Zhouzhi Country	90.42%	Lintong district	82.67%
Beilin District	33.33%	LanTian County	84.93%
Yanta District	55.32%		

**TABLE 6.** Interaction of the communities in the core region.

Community number	Internal interaction ratio	Interaction ratio with other 3 communities	Interaction ratio outside the 4 communities
7	52.01%	32.95%	15.03%
8	54.54%	34.97%	10.49%
9	51.03%	34.73%	14.23%
10	50.74%	39.59%	9.66%

In summary, several official boundaries really respect the human activities, while in some regions, the human cross-border activities are more frequent. The objective division of urban space in this study better reflects the interaction of human beings in urban space, provides a perspective for re-examining the official administrative division, and also provides an opportunity for managers to think about urban social and economic problems. For example, look for the driving factors of human cross-border activities in official regions corresponding to community 3 and community 6 in Figure 10, and then find the hidden urban problems.

## V. CONCLUSION AND FUTURE WORK

In this paper, we have used the cellular data to generate the mobility based spatial interaction network of Xi'an, China, and objectively delineate the urban interior boundaries formed by human activities to re-examine the official administrative division. To ensure capturing the actually regional relevance, we first explore the region-based human crowd mobility patterns at the collective level. The analysis of patterns indicates that region-based human crowd mobility presents some certainty, which is far from random. This conclusion confirms that there is indeed a reliable connection between urban regions due to human mobility. Also, the discovered patterns are used to distinguish anomalous regions, revealing the special structure in the urban. All these patterns and anomalous regions provide the novel perspective to understand the urban dynamics. Then, based on the reliable connection between regions, we objectively divide the closely related regions together. Compared with the official boundaries, several unexpected communities have emerged. These communities really present a closer correlation, respecting to the objective internal interaction ratio, and reveal several naturally differentiated regions within the urban. The obtained divisions offer a perspective to re-examine the official administrative division.

In future work, we will utilize more comprehensive data, including demographic data, regional economic structure

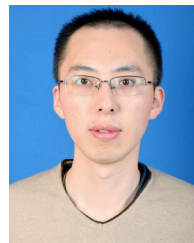
data, land use allocation and other social data, to explore the driving factors of human cross-boundary interaction and provide more detailed arguments for the adjustment of urban administrative divisions.

## REFERENCES

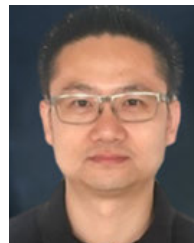
- [1] M. Docu and A. Dunarintu, "The socio-economic impact of urbanization," *Int. J. Academic Res. Accounting, Finance Manage. Sci.*, vol. 2, no. 1, pp. 47–52, 2012.
- [2] Y. Li, Z. Xie, B. Li, and M. Mohiuddin, "The impacts of in situ urbanization on housing, mobility and employment of local residents in China," *Sustainability*, vol. 14, no. 15, p. 9058, Jul. 2022.
- [3] P. Martí, L. Serrano-Estrada, A. Nolasco-Cirugeda, and J. L. Baeza, "Revisiting the spatial definition of neighborhood boundaries: Functional clusters versus administrative neighborhoods," *J. Urban Technol.*, vol. 29, no. 3, pp. 73–94, Jul. 2022.
- [4] G. L. Allen, *Applied Spatial Cognition: From Research to Cognitive Technology*. London, U.K.: Psychology Press, 2020.
- [5] Y. Xu, J. Rollo, D. S. Jones, Y. Esteban, H. Tong, and Q. Mu, "Towards sustainable heritage tourism: A space syntax-based analysis method to improve tourists' spatial cognition in Chinese historic districts," *Buildings*, vol. 10, no. 2, p. 29, Feb. 2020.
- [6] X. Ye and C. Andris, "Spatial social networks in geographic information science," *Int. J. Geograph. Inf. Sci.*, vol. 35, no. 12, pp. 2375–2379, Dec. 2021.
- [7] D. R. Montello, "Cognition of geographic information," in *A Research Agenda for Geographic Information Science*. Boca Raton, FL, USA: CRC Press, 2004, pp. 61–91.
- [8] S. Chang, Z. Wang, D. Mao, F. Liu, L. Lai, and H. Yu, "Identifying urban functional areas in China's Changchun city from Sentinel-2 images and social sensing data," *Remote Sens.*, vol. 13, no. 22, p. 4512, Nov. 2021.
- [9] Y. Chen and A. Deng, "Using POI data and Baidu migration big data to modify nighttime light data to identify urban and rural area," *IEEE Access*, vol. 10, pp. 93513–93524, 2022.
- [10] M. Gandy, "Urban political ecology: A critical reconfiguration," *Prog. Hum. Geogr.*, vol. 46, no. 1, pp. 21–43, Feb. 2022.
- [11] C. Schliephake and U. Schliephake, *Urban Ecologies: City Space, Material Agency, and Environmental Politics in Contemporary Culture*. Lanham, MD, USA: Lexington Books, 2014.
- [12] W. Chen, X. Chen, L. Cheng, X. Liu, and J. Chen, "Delineating borders of urban activity zones with free-floating bike sharing spatial interaction network," *J. Transp. Geogr.*, vol. 104, Oct. 2022, Art. no. 103442.
- [13] M. Jin, L. Gong, Y. Cao, P. Zhang, Y. Gong, and Y. Liu, "Identifying borders of activity spaces and quantifying border effects on intra-urban travel through spatial interaction network," *Comput., Environ. Urban Syst.*, vol. 87, May 2021, Art. no. 101625.



- [14] Y. Shen and M. Batty, "Delineating the perceived functional regions of London from commuting flows," *Environ. Planning A, Economy Space*, vol. 51, no. 3, pp. 547–550, May 2019.
- [15] S. Sobolevsky, M. Szell, R. Campari, T. Couronné, Z. Smoreda, and C. Ratti, "Delineating geographical regions with networks of human interactions in an extensive set of countries," *PLoS ONE*, vol. 8, no. 12, Dec. 2013, Art. no. e81707.
- [16] C. Zhong, S. M. Arisona, X. Huang, M. Batty, and G. Schmitt, "Detecting the dynamics of urban structure through spatial network analysis," *Int. J. Geograph. Inf. Sci.*, vol. 28, no. 11, pp. 2178–2199, Nov. 2014.
- [17] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási, "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, Feb. 2010.
- [18] M. Schlöpfer, L. Dong, K. O'Keeffe, P. Santi, M. Szell, H. Salat, S. Anklesaria, M. Vazifeh, C. Ratti, and G. B. West, "The universal visitation law of human mobility," *Nature*, vol. 593, no. 7860, pp. 522–527, May 2021.
- [19] L. Alessandretti, U. Aslak, and S. Lehmann, "The scales of human mobility," *Nature*, vol. 587, no. 7834, pp. 402–407, Nov. 2020.
- [20] B. Jiang and Y. Miao, "The evolution of natural cities from the perspective of location-based social media," *Prof. Geographer*, vol. 67, no. 2, pp. 295–306, Apr. 2015.
- [21] K. Lynch, *The Image of the City*. Cambridge, MA, USA: MIT Press, 1964.
- [22] H. Barbosa, M. Barthelemy, G. Ghoshal, C. R. James, M. Lenormand, T. Louail, R. Menezes, J. J. Ramasco, F. Simini, and M. Tomasini, "Human mobility: Models and applications," *Phys. Rep.*, vol. 734, pp. 1–74, Mar. 2018.
- [23] S. D. Roches, K. I. Brans, M. R. Lambert, L. R. Rivkin, A. M. Savage, C. J. Schell, C. Correa, L. De Meester, S. E. Diamond, N. B. Grimm, N. C. Harris, L. Govaert, A. P. Hendry, M. T. J. Johnson, J. Munshi-South, E. P. Palkovacs, M. Szulkin, M. C. Urban, B. C. Verrelli, and M. Alberti, "Socio-eco-evolutionary dynamics in cities," *Evol. Appl.*, vol. 14, no. 1, pp. 248–267, 2021.
- [24] Y. Long, H. Han, Y. Tu, and X. Shu, "Evaluating the effectiveness of urban growth boundaries using human mobility and activity records," *Cities*, vol. 46, pp. 76–84, Aug. 2015.
- [25] V. Blondel, G. Krings, and I. Thomas, "Regions and borders of mobile telephony in Belgium and in the Brussels metropolitan zone," *Brussels Stud.*, vol. 42, no. 4, pp. 1–12, 2010.
- [26] Y. Wei, X. Zhao, J. Zou, and E. Herrera-Viedma, "A complementing preference based method for location recommendation with cellular data," *Knowl.-Based Syst.*, vol. 183, Nov. 2019, Art. no. 104889.
- [27] J. Yin, A. Soliman, D. Yin, and S. Wang, "Depicting urban boundaries from a mobility network of spatial interactions: A case study of Great Britain with geo-located Twitter data," *Int. J. Geograph. Inf. Sci.*, vol. 31, no. 7, pp. 1293–1313, Jul. 2017.
- [28] Y. Wei, M. Mao, X. Zhao, J. Zou, and P. An, "City metro network expansion with reinforcement learning," in *Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2020, pp. 2646–2656.
- [29] F. Buono and M. Longobardi, "A dual measure of uncertainty: The Deng entropy," *Entropy*, vol. 22, no. 5, p. 582, May 2020.
- [30] H. Wang, S. Zeng, Y. Li, and D. Jin, "Predictability and prediction of human mobility based on application-collected location data," *IEEE Trans. Mobile Comput.*, vol. 20, no. 7, pp. 2457–2472, Jul. 2021.
- [31] R. Foorthuis, "On the nature and types of anomalies: A review of deviations in data," *Int. J. Data Sci. Anal.*, vol. 12, no. 4, pp. 297–331, Oct. 2021.
- [32] S. Thudumu, P. Branch, J. Jin, and J. Singh, "A comprehensive survey of anomaly detection techniques for high dimensional big data," *J. Big Data*, vol. 7, no. 1, pp. 1–30, Dec. 2020.
- [33] M. Zhang, T. Li, Y. Yu, Y. Li, P. Hui, and Y. Zheng, "Urban anomaly analytics: Description, detection, and prediction," *IEEE Trans. Big Data*, vol. 8, no. 3, pp. 809–826, Jun. 2022.
- [34] K. Kalair and C. Connaughton, "Anomaly detection and classification in traffic flow data from fluctuations in the flow–density relationship," *Transp. Res. C, Emerg. Technol.*, vol. 127, Jun. 2021, Art. no. 103178.
- [35] M. Zhang, T. Li, H. Shi, Y. Li, and P. Hui, "A decomposition approach for urban anomaly detection across spatiotemporal data," in *Proc. IJCAI*, Aug. 2019, pp. 6043–6049.
- [36] E. C. M. Hui, X. Li, T. Chen, and W. Lang, "Deciphering the spatial structure of China's megacity region: A new bay area—The Guangdong-Hong Kong-Macao Greater Bay area in the making," *Cities*, vol. 105, Oct. 2020, Art. no. 102168.
- [37] M. Roosta, M. Javadpoor, and M. Ebadi, "A study on street network resilience in urban areas by urban network analysis: Comparative study of old, new and middle fabrics in Shiraz," *Int. J. Urban Sci.*, vol. 26, no. 2, pp. 309–331, Apr. 2022.
- [38] F. T. Liu, M. T. Kai, and Z. H. Zhou, "Isolation forest," in *Proc. 8th IEEE Int. Conf. Data Mining*, 2008, pp. 413–422.
- [39] M. Curado, L. Tortosa, J. F. Vicent, and G. Yeghikyan, "Analysis and comparison of centrality measures applied to urban networks with data," *J. Comput. Sci.*, vol. 43, May 2020, Art. no. 101127.
- [40] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *J. Stat. Mech., Theory Exp.*, vol. 2008, no. 10, Oct. 2008, Art. no. P10008.
- [41] M. E. J. Newman, "Modularity and community structure in networks," *Proc. Nat. Acad. Sci. USA*, vol. 103, no. 23, pp. 8577–8582, Jun. 2006.
- [42] S. Fortunato, "Community detection in graphs," *Phys. Rep.*, vol. 486, nos. 3–5, pp. 75–174, Feb. 2010.
- [43] M. Rosvall and C. T. Bergstrom, "Maps of random walks on complex networks reveal community structure," *Proc. Nat. Acad. Sci. USA*, vol. 105, no. 4, pp. 1118–1123, Jan. 2008.
- [44] P. Pons and M. Latapy, "Computing communities in large networks using random walks," *J. Graph Algorithms Appl.*, vol. 10, no. 2, pp. 191–218, 2006.



**YU WEI** was born in 1990. Since January 2021, he has been a University Teacher with Taiyuan University of Technology. His research interests include urban planning and machine learning.



**XI ZHAO** received the B.S. degree in information engineering and the M.S. degree in systems engineering from Xi'an Jiaotong University, in 2003 and 2007, respectively, and the Ph.D. degree in information and data from the University of Central Technology, France, in 2010. He was a Postdoctoral Researcher at the Department of Computer Science, University of Houston, USA, and he was promoted as a Research Assistant Professor. In 2015, he was selected as a Special Researcher of the Top Notch Talent Program, Xi'an Jiaotong University. In 2018, he was selected as a Professor of the Top Notch Talent Program, Xi'an Jiaotong University.

...