

Research Article

Using Smart Card Data of Metro Passengers to Unveil the Urban Spatial Structure: A Case Study of Xi'an, China

Guohong Cheng ¹, Shichao Sun ², Linlin Zhou ² and Guanzhong Wu ³

¹Hangzhou Communications Planning and Design Institute Co., Ltd, Hangzhou, China

²College of Transportation Engineering, Dalian Maritime University, Dalian, China

³Shanghai Tongzhun Information Technology Co., Ltd, Shanghai, China

Correspondence should be addressed to Shichao Sun; sunshichao1988@hotmail.com

Received 3 October 2021; Accepted 19 October 2021; Published 3 November 2021

Academic Editor: Jinyu Chen

Copyright © 2021 Guohong Cheng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study adopted smart card data collected from metro systems to identify city centers and illustrate how city centers interacted with other regions. A case study of Xi'an, China, was given. Specifically, inflow and outflow patterns of metro passengers were characterized to measure the degree of population agglomeration of an area, i.e., the centrality of an area. On this basis, in order to overcome the problem of determining the boundaries of the city centers, Moran's I was adopted to examine the spatial correlation between the inflow and outflow of ridership of adjacent areas. Three residential centers and two employee centers were identified, which demonstrated the polycentricity of urban structure of Xi'an. With the identified polycenters, the dominant spatial connections with each city center were investigated through a multiple linkage analysis method. The results indicated that there were significant connections between residential centers and employee centers. Moreover, metro passengers (commuters mostly) flowing into the identified employee centers during morning peak-hours mainly came from the northern and western area of Xi'an. This was consistent with the interpretation of current urban planning, which validated the effectiveness of the proposed methods. Policy implications were provided for the transport sector and public transport operators.

1. Introduction

Uncovering city structure and urban spatial connections can benefit from appropriate allocation of certain kinds of resources, e.g., land use, medical facilities, educational resources, and transport infrastructures. Specifically, the underlying urban structure is commonly interpreted beyond its physical form, because cities run as dynamic systems [1]. Therefore, the urban spatial connections beneath the complex system are not only related to the distribution of physical environments and economic resources, but also heavily involved with intracity movements [2,3]. Flows of people or cargo function as ties that integrate static physical resources into a dynamic system and generate spatial interactions [4–7]. In this context, a large number of studies have unveiled the city structure via passenger flow systems [8–10]. Specifically, the centers within city structure are found highly related to the spatial agglomeration of

population [8], while the flow patterns of passengers are commonly utilized to denote how the identified city centers interact with other regions [11]. Given these practical enlightenments, policy implications and targeted strategies are always developed to guide the urban planning and hinder uncontrolled urban sprawl [12].

It has drawn much attention of researchers to investigate the urban spatial structure by using trajectory-based data [13], such as taxi GPS data [1,5], smart card data [8,12], and social network check-in data [9]. In this respect, Liu et al. [1] studied passengers' travel patterns and detected the polycentric city structure of Shanghai, China, by using taxi trajectory data. Tanahashi et al. [14] applied graph-partitioning methods to analyze the flow patterns of people between partitioned subregions by the use of mobile phone data and then revealed the urban spatial structure associated with travel activities. Yu et al. [15] proposed a methodology framework involving the community detection method and

mobile phone data, through which the urban structure was described by decomposing commuting demands. However, taxi trajectory data are plagued by the population coverage, and mobile phone data involve privacy and security issues.

In addition to taxi trajectory data and mobile phone data, smart card data are also recognized as a promising data source to provide insights on the identification of urban spatial structure [16,17]. It is because smart card data can provide rich and high-quality check-in records of public transport passengers [18–22], and mostly important, these passengers' riderships constitute a crucial part of urban spatial movements [21, 23, 24]. Compared with other data sources, smart card data are accessible with less cost, and the data are refined in spatial and temporal granularity [25]. In addition, the coverage of smart card data is relatively wide both in space and in population [23]. Tang et al. [26] proposed a clustering refinement approach to investigate the agglomeration pattern of passenger flows by using smart card data and then elaborated five clusters of metro stations to represent the underlying structure. Long and Shen [27] combined smart card data and POIs (points of interest) and proposed a clustering approach to identify different functional zones of the city and understand their spatial distribution characteristics. Nevertheless, the two studies did not conduct an in-depth analysis of urban spatial connections. Gong, Lin and Duan [4] used principle component analysis method to decompose the passenger flows obtained from the Automatic fare collection (AFC) data and investigated the spatiotemporal structure of urban form. Zhong et al. [28] used the smart card data collected from different time periods in Singapore and then investigated the overall spatial structure by monitoring urban movements from daily transportation. However, compared with the studies of [26, 27], they paid more attention to the spatial connections of discrete regions, but little was discussed on the identification of the city centers within the urban structure.

With respect to the methods used to reveal urban spatial connections, a number of previous studies adopted community detection technology to find the substructures of a complex spatial network [29]. However, the communities detected in the urban spatial network only denote the regions where their internal spatial connections are obviously stronger than those interacting with other areas [1]. Thus, it is weak to employ the community detection method to investigate the spatial connections between city centers and other discrete regions. Graph-partitioning method and spatial clustering approach are another two conventional methods commonly used to identify city centers and then reveal urban spatial interactions. Compared with community detection technology, these two methods can better illustrate the polycenters and regional connections, which constitute the main city structure [4, 8, 14]. Nevertheless, it is still not a straightforward task to determine the boundaries of city centers, which are regarded as key nodes in the spatial connections. For instance, Roth et al. [8] employed smart card data collected from London Underground to identify the urban spatial structure through a spatial clustering approach. However, the data adopted in their study were collected at station-based level, so that the basic O-D matrix obtained to illustrate the spatial connections and could

only be described at station-to-station granularity. Thus, this led to the problem that it was difficult to determine which adjacent stations should be merged and then employed to represent the regions of city centers. In this context, the urban spatial structure could not be further interpreted in a more specific, refined and microperspective way.

Therefore, this study tried to use the smart card data collected from a metro system to reveal the urban spatial structure, through a spatial analysis approach. Compared with previous studies, the contributions of this paper could be summarized as follows:

- (i) In order to identify the city centers and precisely determine their boundaries, the spatial autocorrelation analysis method was employed to merge the adjacent regions where the characteristics of trip generation were highly correlated.
- (ii) Based on the O-D matrix obtained from smart card data, the multiple linkage analysis method was proposed to illustrate how the identified city centers dynamically interacted with other discrete regions, resulting in the dominated structure of regional connections.

The remainder of this paper is organized as follows. Section 2 gives a description of the study area and datasets involved in this paper. Section 3 introduces the methods employed in this study. The results and corresponding discussions are elaborated in Section 4. Conclusions and future works are drawn in Section 5.

2. Study Area and Datasets

2.1. Study Area. Xi'an city is the capital of Shaanxi province in north-central China. According to the census in 2019, the city is administering 11 districts and 2 counties with nearly 10-million population. Xi'an Metro refers to the rail transit system serving the urban area of the city. It was first open for operation in September, 2011, and up to 2021, it has eight metro lines (153 stations) with a total length of 244 kilometers. However, the smart card data used for this study were harvested in 2018. Therefore, this study mainly focused on the approach proposed to reveal the urban spatial structure but took the relatively outdated urban form of Xi'an as a demonstration. In this respect, regardless of the dynamic changes of urban form, the proposed approach can retain its resilience, which can be justified through the following empirical study. Besides, only three metro lines had been put into operation in 2018. Nevertheless, as the capital of ancient China, the urban form of Xi'an has always been an axial development. Therefore, the first three metro lines basically covered the main functional areas of the city. That is, the smart card data collected from the metro system in 2018 could be used to demonstrate the dominated urban spatial structure at that time. In addition, from the perspective of ridership, the daily average passenger ridership of Xi'an Metro had exceeded 2 million, which accounted for 30.3% of the total [30]. Thus, its coverage of passenger flows was wide enough and even better than taxi trajectory data used in previous studies, to unveil the overall spatial interactions.

2.2. Datasets. Up to 2018, over 85% of transactions in Xi'an public transit system were completed through smart cards. This proportion would reach nearly 100% in the metro (fare evasion behavior was not considered). Thus, it implied that smart card data collected from Xi'an Metro would record all the transaction information. The smart card data used in this study were collected from the AFC systems of Xi'an Metro. The data were not linked with passengers' bus trips. Thus, only the trips generated in the metro network were included. The dataset covered nearly 10 million transaction records, which were generated during the period from 17 April, 2018, to 21 April, 2018. The raw data were preprocessed by service providers, and invalid data were filtered out, which contained incomplete travel information. With each record of the data, we could obtain the card number (unique for each cardholder), the transaction time, the inbound information, the outbound information, and the transaction amount, as shown in Table 1.

3. Methods

3.1. The Conceptual Framework. The high agglomeration of population has been considered as good proxies for evaluating the centrality of an area [8]. However, the population distribution in a city is dynamically changing, which results from people's daily intracity movements. In this context, the centrality degree of an area shall be assessed accordingly by its inflow as well as outflow of public transport ridership. In addition, passenger flows within a day are commonly back and forth; e.g., a typical commuting can be characterized as going to work in the morning and returning home after work. Thus, it may lead to the fact that a potential city center, which attracts a large inflow of ridership, may also have almost equal outflow of ridership within a day. In this respect, we categorized the city centers into residential centers and employee centers. A residential center commonly refers to an area where the living density is very high, and a large outflow of passengers are generated in the morning due to the commuting demand, while an employee center can be defined as an area where commercial and industrial activities are frequent, and a large inflow of passengers are attracted. In order to distinguish between residential centers and employee centers, we assumed that the outflow of passengers of residential centers during morning peak-hours should be obviously larger than other areas. In contrast, the characteristic of passenger flows of an employee center would be exactly the opposite. That is, the inflow of passengers of employee centers should dominate in the morning.

Specifically, the outflow of passengers of an area could be measured by the inbound ridership of corresponding metro stations. Similarly, the outbound ridership of metro stations could be used to denote the inflow of passengers of the area where these stations served. However, a city center may cover a larger area served by multiple metro stations. Thus, in order to identify the city centers and their boundaries, it is necessary to investigate not only the inflow and outflow of passengers of metro stations, but also the spatial correlation between them. With respect to this, we tried to use global and local Moran's I to achieve the spatial correlation analysis

between inflow or outflow of ridership of adjacent metro stations. Then, the identified hotspot areas and a part of outliers could be regarded as the city centers. On this basis, we attempted to examine the interactions between city centers and other regions, so as to illustrate the urban spatial structure of Xi'an. In particular, we provided insights on the forms in which passenger flows were distributed based on the O-D matrix extracted from the smart card data. The multiple linkage analysis method was employed to determine the significant connections of each city center.

3.2. Global Moran's I. Spatial autocorrelation analysis has been widely used in GIS to better understand the spatial dependency between one object with other nearby objects. In other words, it was commonly used to measure the degree to which one spatial area is similar to other vicinities. Spatial autocorrelation is multidimensional, thus being more complex than conventional one-dimensional autocorrelation. In 1950s, Moran [31] firstly developed Moran's I (Index) to measure spatial autocorrelation based on both feature locations and feature values simultaneously. Global Moran's I is defined as follows:

$$I = \frac{N}{S_0} \frac{\sum_i \sum_j w_{ij} (X_i - \bar{X})(X_j - \bar{X})}{\sum_i (X_i - \bar{X})^2}, \quad (1)$$

where spatial objects are indexed by i and j , and the number of which is denoted as N ; X_i represents the vector of feature values of object i ; \bar{X} is the mean of X_i ; w_{ij} is the matrix of spatial weights between objects i and j , and the diagonal of the matrix is with zeros; S_0 is the sum of all w_{ij} . In addition, w_{ij} is found to exert a strong influence on the value of Moran's I, and the distance decay function is commonly used to assign the spatial weights.

Based on the value of Moran's I (range from -1 to 1), we can classify the spatial dependency between the objects in space as positive, negative, and no correlation. Among them, positive spatial dependency (the value of Moran's I exceeds 0) implies that feature values of the objects are similar and clustered together in space. On the contrary, negative spatial dependency (the value of Moran's I is less than 0) is obtained when similar feature values of objects are dispersed in space. Regarding no spatial correlation, it indicates that the spatial distribution of objects' features is random. Thus, spatial autocorrelation analysis can be used to indicate whether there is clustering or dispersion in space, e.g., city centers where populations are spatially agglomerated. In addition, the statistical significance of Moran's I Index is commonly assessed by Z-score as well as P-value. Specifically, Z-score is suggested to be greater than 1.96 or smaller than -1.96, which can demonstrate positive or negative spatial dependency at the 5% significance level.

3.3. Anselin Local Moran's I. Even, given Moran's I, it still cannot help us directly identify statistically significant hotspot areas (i.e., city centers to be determined in this paper), cold spot areas, and spatial outliers in space. As an

extension of global Moran's I, Anselin local Moran's I was developed by Anselin [32], and it was defined as follows:

$$I_i = \frac{X_i - \bar{X}}{S_i^2} \sum_{j=1, j \neq i}^N w_{ij} (X_j - \bar{X}), \quad (2)$$

$$S_i^2 = \frac{\sum_{j=1, j \neq i}^N (X_j - \bar{X})^2}{N - 1},$$

where N denotes the number of spatial objects, which are indexed by i and j ; X_i represents the vector of feature values of object i ; \bar{X} is the mean of X_i ; w_{ij} is the spatial weight between objects i and j .

Other than Z-score and P-value, which are used to assess the statistical significance of local Moran's I, clustering or outlier types will be attached to each study object. Specifically, the values of Z-score and Lisa are simultaneously positive or negative when the study object is surrounded by other objects with similar values in space. Thus, it demonstrates a typical partial clustering in space, and the type L-L or H-H can be attached to the clustering. Clustering type H-H indicates a high-feature-value clustering, while clustering type L-L implies the clustering consisting of objects with low feature value. We can distinguish the clustering type H-H from the clustering type L-L by evaluating the value of Lisa (clustering type H-H can be determined when the value of Lisa is positive; otherwise, clustering type L-L will be attached). On the other side, a study object can be regarded as an outlier in space when the value of Z-score and the value of Lisa have different signs. It demonstrates that the study object is surrounded by other objects with dissimilar values. Specifically, the outlier will be attached with clustering type H-L when it owns high feature values but surrounded by others with low feature values. Conversely, the outlier with clustering type L-H is characterized by low feature values but surrounded by others with

high feature values. Thus, as defined with the clustering type H-H and the outlier type H-L, the city centers where large populations are agglomerated in partial spatial areas can be identified by adopting local Moran's I, because it can be used to reflect the characteristics of spatial clustering of ridership.

3.4. Multiple Linkage Analysis Based on O-D Matrix. It has been long recognized as a straightforward task to obtain the O-D matrix by using smart card data collected from metro systems [18, 19, 33]. Specifically, the inbound information can provide the passengers' origin stations, while the outbound information can be used to infer the destination stations. In this context, the O-D matrix derived from the smart card data can be utilized to reflect passenger flows and underlying travel patterns in both spatial and temporal dimensions. Thus, smart card data own the natural property to illustrate the dynamic urban spatial interactions related to the movements of people.

Nevertheless, it is still necessary to find dominant passenger flows that construct the main city structure, because most of the elements involved in the O-D matrix are not significant to reflect the primary spatial connections of a city. Therefore, in order to distinguish between significant flows and insignificant flows, this paper employed multiple linkage analysis method to investigate dominant passenger flows based on the O-D matrix [12]. Regarding multiple linkage analysis, it assumes that there are a total of $k + 1$ centers within the spatial structure. All flows from or to each spatial center are sorted from the largest (W_1) to the smallest (W_k) by their passenger ridership. These ordered flows constitute a set of observed flows $\{W_k\}$. Moreover, a set of expected flows $\{W_k'\}$ is generated for the same spatial center by different cycles, and the definition is as follows:

$$1^{st} \text{ cycle: } W'_1 = \sum_{i=1}^K W_i, \quad W'_2 = W'_3 = W'_4 \cdots = W'_k = 0,$$

$$2^{nd} \text{ cycle: } W'_1 = W'_2 = \frac{1}{2} \sum_{i=1}^K W_i, \quad W'_3 = W'_4 \cdots = W'_k = 0,$$

$$j^{th} \text{ cycle } (j < k): W'_1 = W'_2 = \cdots = W'_j = \frac{1}{j} \sum_{i=1}^K W_i, \quad W'_{j+1} = W'_{j+2} \cdots = W'_k = 0,$$

$$k^{th} \text{ cycle: } W'_1 = W'_2 = \cdots = W'_k = \frac{1}{k} \sum_{i=1}^K W_i.$$

The goodness-of-fit between the set of observed flows and each cycle of expected flows is measured by R-square [12]. Then, the number of dominated flows from or to each center can be determined by finding the j^{th} cycle where the highest

R-square value places. In short, the mechanism of multiple linkage analysis is to minimize the difference between the real configurations and ideal configurations where the flows are distributed over each link in shares of equal magnitude.

TABLE 1: Smart card datasets.

Card #	Date	Time	In-station	In-line	Out-station	Out-line	Amount
51050***	20180420	06:12:51	Beiyuan	2	Beikezhan	2	2 RMB
68800***	20180420	06:13:21	Banpo	1	Fangzhicheng	1	2 RMB
50096***	20180420	06:16:42	Zaohe	1	Houweizhai	1	2 RMB

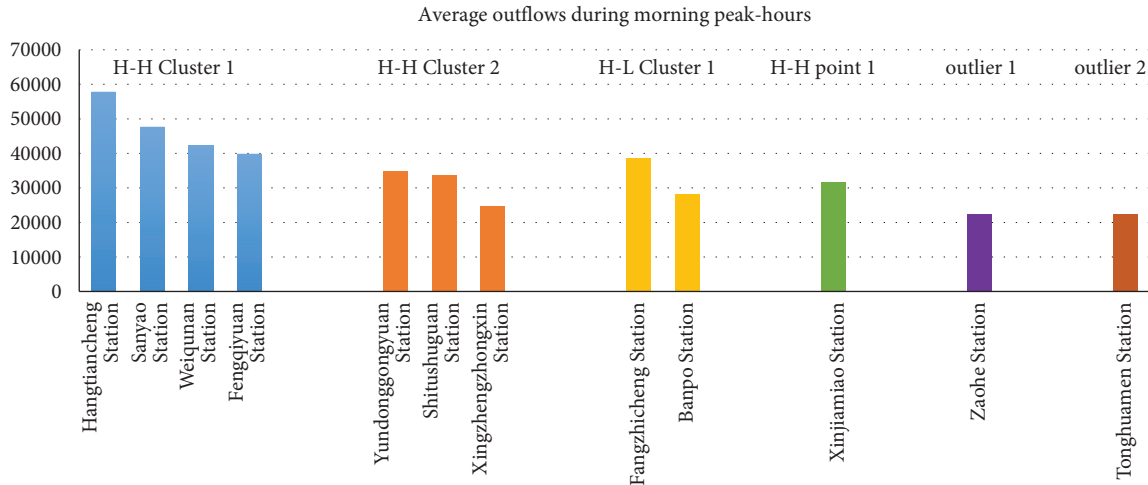
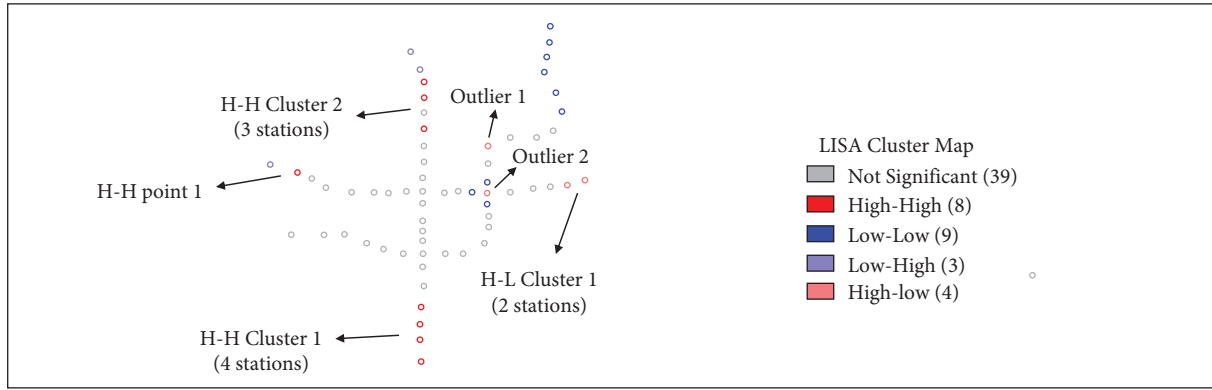


FIGURE 1: Spatial autocorrelation analysis on the outflow of ridership of metro stations.

4. Results and Discussions

4.1. Determination of the Polycenters of Urban Structure.

In this study, global Moran’s I was used to measure whether there was a spatial correlation of ridership (including inflow and outflow) between adjacent metro stations. Then, local Moran’s I was further adopted to identify where the cluster or outlier was. Specifically, the morning peak-hours of Xi’an on weekdays was set as 06:00 am to 09:00 am. The average inflow and outflow of ridership of each metro station during morning peak-hours were obtained from the smart card data. The software GeoDa (version 1.12) was used to analyze the spatial dependency in the map, where an embedded default distance decay function was employed to build the spatial weights matrix. The results were shown as follows.

First, the global Moran’s I regarding the outflow of ridership was 0.2881 with a Z-score of 5.1169, greater than the cut-off value (1.96) at 0.05 significance level. Thus, it demonstrated a positive spatial autocorrelation, namely,

spatial clustering of passenger outflows in the map. On this basis, given the Lisa cluster map shown in Figure 1, it indicated that two clusters were identified with type H–H, namely, H–H Cluster 1 (consisted of 4 stations) and H–H Cluster 2 (consisted of 3 stations). Moreover, a spatial agglomeration of H–L outliers was identified, called H–L Cluster 1 (consisted of 2 stations). The results indicated that the outflow of ridership of the above identified three clusters was dominant, which averagely accounted for more than 35% of the daily total during morning peak hours on weekdays. Thus, combined with further validation through the current status data of land use, these three clusters could be recognized as residential centers. Other than spatial clusters, a H–H point (namely, Xinjiamiao Station) and two H–L points (namely, Zaohe Station and Tonghuamen Station) were identified. Nevertheless, they were all relatively isolated in space with no significant passenger outflows. In summary, three residential centers were finally determined, and we, respectively, named H–H Cluster 1 as

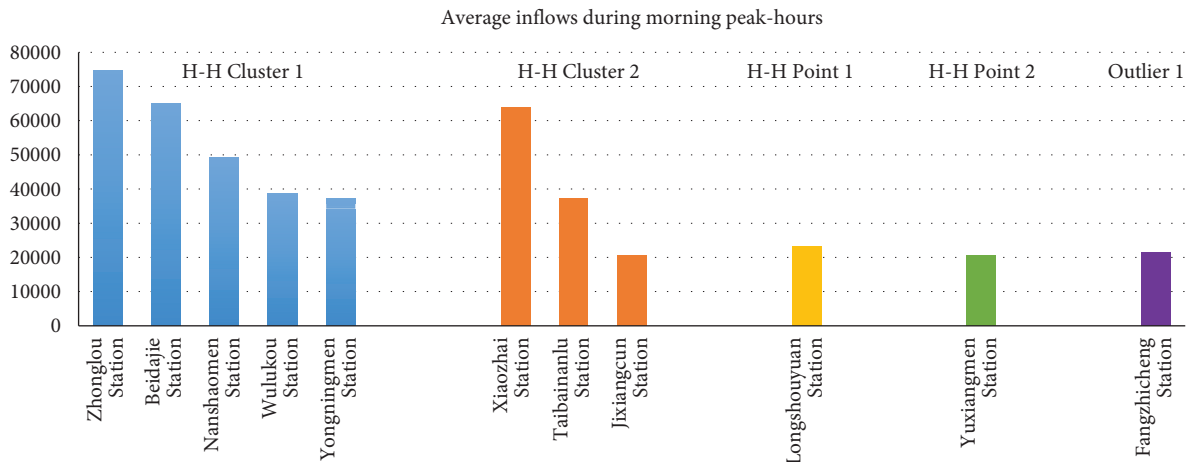
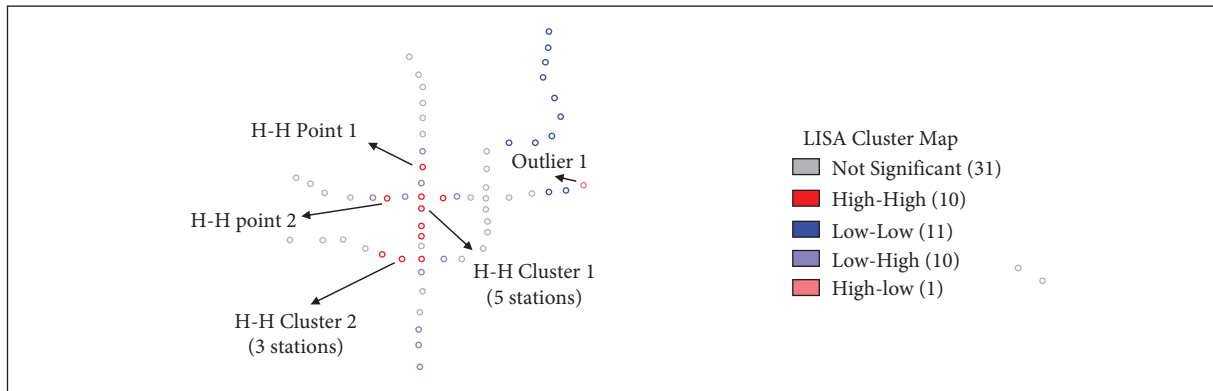


FIGURE 2: Spatial autocorrelation analysis on the inflow of ridership of metro stations.

Hangtiancheng Center, H-H Cluster 2 as Yundonggongyuan Center, and H-L Center as Fangzhicheng Center.

On the other side, the results of spatial autocorrelation analysis on the inflow of ridership are given in Figure 2. They indicated that the value of global Moran's I was 0.3472 with Z -score (5.5868) greater than the cut-off value (1.96) at the 0.05 significance level. Thus, it implied that the inflow of ridership was also spatially clustered in the map. On the basis of Lisa cluster map shown in Figure 2, two H-H clusters were identified, which were, respectively, defined as H-H Cluster 1 (consisted of 5 stations) and H-H Cluster 2 (consisted of 3 stations). The inflows of these two clusters accounted for more than 40% of the daily total during morning peak hours on weekdays. In addition to the clusters, two H-H points and one H-L point were identified, but regarded as isolated outliers with no significant inflow of ridership. Therefore, two employee centers were finally determined, and we, respectively, named H-H Cluster 1 as Zhonglou Center and H-H Cluster 2 as Xiaozhai Center.

In summary, based on the results of above spatial autocorrelation analysis, we identified three residential centers and two employee centers. The residential centers included Hangtiancheng Center (4 stations), Yundonggongyuan Center (3 stations), and Fangzhicheng Center (2 stations), while the employee centers consisted of Zhonglou Center (5 stations) and Xiaozhai Center (3 stations). Therefore, the results demonstrated the polycentricity of urban structure of Xi'an.

4.2. Determination of the Significant Flows of Polycenters.

With the identified polycenters, the significant flows connecting with each city center were investigated. Specifically, MLA was applied to the O-D matrix between identified clusters and other metro stations, which was obtained from the smart card data. The number of significant flows of each city center is listed in Table 2. The results indicated that the number of significant connections of employee centers was obviously larger than that of residential centers. Therefore, it implied that the inflows of passengers of employee centers were more evenly distributed on the dominant linkages. On the contrary, the significant outflows of passengers of residential centers were relatively concentrated on less linkages.

4.3. Determination of the Urban Spatial Structure of Xi'an.

In order to intuitively illustrate how the identified polycenters interacted with other regions, the urban spatial structure of Xi'an was revealed by demonstrating the significant linkages of each city center through the software NodeXL, as shown in Figure 3.

Specifically, the results indicated that there existed strong connections between residential centers and employee centers. For instance, all the residential centers have significant connections with Zhonglou Center, which was identified as one of the employee centers in Xi'an. In addition, Yundonggongyuan Center and Hangtiancheng

TABLE 2: The number of significant flows of each city center

Center type	Center	The number of significant flows
Residential center	Hangtiancheng center	12
	Yundonggongyuan center	6
	Fangzhicheng center	5
Employee center	Zhonglou center	20
	Xiaozhai center	14

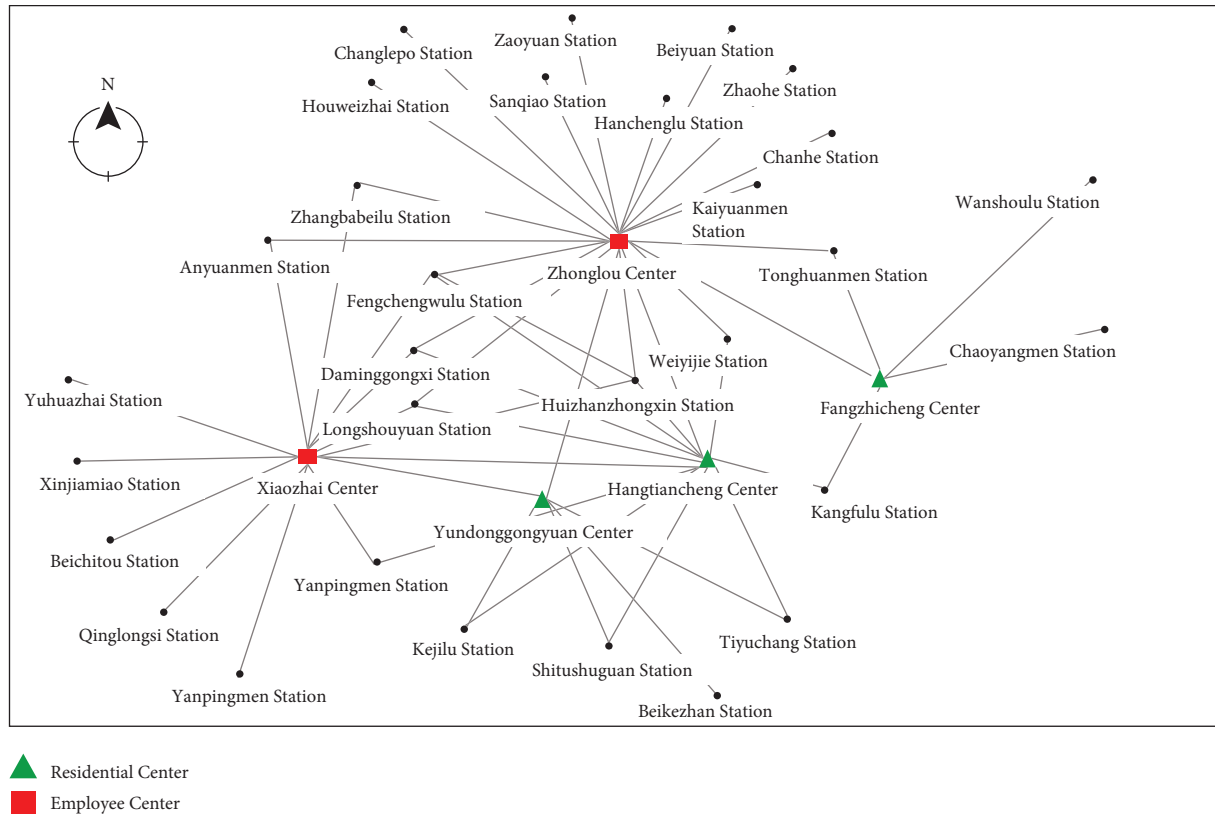


FIGURE 3: The urban spatial structure of Xi'an.

Center, two of the identified residential centers, were found to have strong connections with Xiaozhai Center, the other employee center. It also showed that the ridership from the identified residential centers to the employee centers accounted for nearly 40% of the total significant flows. On the other side, the spatial interactions reflected by passenger flows between the residential centers or the employee centers were not significant. Thus, the urban spatial structure could be further divided into two hierarchies, i.e., a fundamental spatial structure and a comprehensive spatial structure. The comprehensive spatial structure is shown in Figure 3, while the spatial connectivity between the identified residential and employee centers could be regarded as a fundamental spatial structure beyond the comprehensive one.

4.4. Discussions. Other than the urban spatial structure, commuting demand by metro could be reflected by the identified dominated flows of passengers during morning peak-hours. Specifically, as shown in Figure 4, it illustrated

the dominant inflows of ridership of Zhonglou Center; the layout had been adjusted according to the actual geography.

The results indicated that metro passengers (commuters mostly) flowing into Zhonglou Center in morning peak-hours mainly came from the northern and western area of Xi'an, which accounted for more than 70% of the total inflows of the center. It was consistent with the current urban planning and land use situation of Xi'an. Specifically, Zhonglou Center refers to the area around the Xi'an Bell Tower, which is located in the geographical center of the city; at the same time, the earliest development of commerce of Xi'an also began in the Bell Tower area. Thus, it was not a surprise to see the Bell Tower area to be identified as an employee center. In addition, the northern and western regions of Xi'an were developed earlier, and according to the overall urban planning of Xi'an, these regions were planned to be developed in the form of small but scattered areas. Thus, this was why most of the dominated inflows were concentrated in these two regions but scattered on multiple linkages. Separately, the identified three residential centers

dispersed but mainly came from the north area of Xi'an. Thus, it was a surprise to find that although Xiaozhai Center was located in the south of the city, there was still a lot of commuting demand from the north. Actually, according to the development of urban planning of Xi'an, Xiaozhai Center could be regarded an emerging urban commercial center. Nevertheless, since the northern regions of the city were developed earlier with very completed functions, most populations preferred to live in the north of the city. Therefore, it led to the large long-distance commutes from the north to the south.

Overall, the urban spatial structure of Xi'an unveiled in this study was reasonable, including the city centers identified by spatial correlation analysis and dominant spatial connections obtained by MLA. Thus, to some extent, this recognized the effectiveness of the methods presented in this study. As for policy implications, the interpretation of the structure of passenger flows would benefit the transport sector and public transport operators to determine the development direction. For example, nonstop service, express service, and all-stop service can be mixed and organized for Xi'an Metro. Specifically, nonstop service can be implemented to smooth the connectivity between residential centers and employee centers. Express service can be adopted for the direction of dominant spatial connections during morning peak-hours.

5. Conclusions

As probably one of the most complex spatial systems, the city structure is highly correlated with urban movements of people or cargo. Thus, the urban spatial structure is not static but should be interpreted from the perspective of dynamic flows. The understanding of the underlying urban structure can contribute to targeted strategies that can better guide the overall urban planning and hinder uncontrolled sprawl in urban areas. This paper used the characteristics of metro passenger flows, which were obtained from the smart card data, to reveal the urban spatial structure. A case study of Xi'an, China, was given, and conclusions were drawn as follows:

- (i) Spatial correlations of inflows and outflows of ridership of metro stations during weekday morning peak-hours were examined through global and local Moran's I. The clustering of spatially adjacent metro stations, which had a high agglomeration of passengers, was identified as city centers. The identified city centers were subdivided into residential centers and employee centers based on the flow direction of passengers. As a consequence, three residential centers and two employee centers were determined, which demonstrated the polycentricity of urban structure of Xi'an.
- (ii) In order to investigate the dominant flows of the identified city centers, the method of MLA was proposed to unveil the most significant spatial connections that constituted the city structure. A two-hierarchical structure was revealed, including a

fundamental spatial structure, which consisted of the strong spatial interactions between the residential centers and the employee centers, and a comprehensive spatial structure, which consisted of all the dominant connections with the identified city centers.

- (iii) Dominant inflows of ridership of employee centers during morning peak-hours were illustrated, and they were interpreted based on the overall urban planning of Xi'an. The results indicated that the identified urban spatial structure was reasonable, which recognized the effectiveness of the methods proposed in this study. In addition, policy implications were provided for the transport sector and public transport operators. For example, nonstop service, express service, and all-stop service can be mixed and organized for Xi'an Metro.

However, this study also has its limitations that should be improved in future researches. First, compared with the metro network of Shanghai, Beijing, and Guangzhou, the network scale of Xi'an Metro is relatively small. It leads to huge differences between them in terms of passenger flow and rail transit coverage. Thus, the adaptability of the methods proposed in this study still needs further verification by using the data collected from other large-scale metro systems. Second, due to the emergence of bike-sharing or MASS systems, there are more and more transport modes that can be used to connect the metro system. In this context, other than the O-D matrix obtained from the metro AFC system, the actual trip distribution of metro passengers should also be brought into future studies.

Data Availability

Access to data is restricted because of third-party rights and personal privacy.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by the Open Funding Project of Key Laboratory of Road and Traffic Engineering of the Ministry of Education (Tongji University) and the Fundamental Research Funds for the Central Universities (3132019163).

References

- [1] X. Liu, L. Gong, Y. Gong, and Y. Liu, "Revealing travel patterns and city structure with taxi trip data," *Journal of Transport Geography*, vol. 43, pp. 78–90, 2015.
- [2] Y. Yue, H.-d. Wang, B. Hu, Q.-Q. Li, Y.-G. Li, and A. G. O. Yeh, "Exploratory calibration of a spatial interaction model using taxi GPS trajectories," *Computers, Environment and Urban Systems*, vol. 36, no. 2, pp. 140–153, 2012.

- [3] Y. Song, L. Merlin, and D. Rodriguez, "Comparing measures of urban land use mix," *Computers, Environment and Urban Systems*, vol. 42, pp. 1–13, 2013.
- [4] Y. Gong, Y. Lin, and Z. Duan, "Exploring the spatiotemporal structure of dynamic urban space using metro smart card records," *Computers, Environment and Urban Systems*, vol. 64, pp. 169–183, 2017.
- [5] Y. Liu, F. Wang, Y. Xiao, and S. Gao, "Urban land uses and traffic 'source-sink areas': evidence from GPS-enabled taxi data in Shanghai," *Landscape and Urban Planning*, vol. 106, no. 1, pp. 73–87, 2012.
- [6] L. Hu, J. Yang, T. Yang, Y. Tu, and J. Zhu, "Urban spatial structure and travel in China," *Journal of Planning Literature*, vol. 35, no. 1, pp. 6–24, 2020.
- [7] Y. Long and J.-C. Thill, "Combining smart card data and household travel survey to analyze jobs-housing relationships in Beijing," *Computers, Environment and Urban Systems*, vol. 53, pp. 19–35, 2015.
- [8] C. Roth, S. M. Kang, M. Batty, and M. Barthelemy, "Structure of urban movements: polycentric activity and entangled hierarchical flows," *PLoS One*, vol. 6, 2011.
- [9] Y. Liu, Z. Sui, C. Kang, and Y. Gao, "Uncovering patterns of inter-urban trip and spatial interaction from social media check-in data," *Plos One*, vol. 9, Article ID e86026, 2014.
- [10] N. Green, "Functional polycentricity: a formal definition in terms of social network analysis," *Urban Studies*, vol. 44, no. 11, pp. 2077–2103, 2007.
- [11] C. Kang, X. Ma, D. Tong, and Y. Liu, "Intra-urban human mobility patterns: an urban morphology perspective," *Physica A: Statistical Mechanics and Its Applications*, vol. 391, no. 4, pp. 1702–1717, 2012.
- [12] S. Sun, Z. Duan, D. Yang, and W. Li, *Polycentricity of the Urban Structure: Spatial Movements Analysis in Shanghai with Smart Card Data*, World Congress Conference, Detroit, Michigan, 2014.
- [13] Y. Yue, T. Lan, A. G. O. Yeh, and Q.-Q. Li, "Zooming into individuals to understand the collective: a review of trajectory-based travel behaviour studies," *Travel Behaviour And Society*, vol. 1, no. 2, pp. 69–78, 2014.
- [14] Y. Tanahashi, J. R. Rowland, S. North, and K. L. Ma, "Inferring human mobility patterns from anonymized mobile communication usage," in *Proceedings of the 10th International Conference on Advances in Mobile Computing and Multimedia*, pp. 151–160, New York, NY, USA, July 2012.
- [15] Q. Yu, W. F. Li, D. Y. Yang, and H. R. Zhang, "Mobile phone data in urban commuting: a network community detection-based framework to unveil the spatial structure of commuting demand," *Journal of Advanced Transportation*, vol. 2020, p. 15, 2020.
- [16] M.-P. Pelletier, M. Trépanier, and C. Morency, "Smart card data use in public transit: a literature review," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 4, pp. 557–568, 2011.
- [17] Y. Sun, J. Shi, and P. M. Schonfeld, "Identifying passenger flow characteristics and evaluating travel time reliability by visualizing AFC data: a case study of Shanghai Metro," *Public Transport*, vol. 8, no. 3, pp. 341–363, 2016.
- [18] X. Ma, C. Liu, H. Wen, Y. Wang, and Y.-J. Wu, "Understanding commuting patterns using transit smart card data," *Journal of Transport Geography*, vol. 58, pp. 135–145, 2017.
- [19] X. Ma, Y.-J. Wu, Y. Wang, F. Chen, and J. Liu, "Mining smart card data for transit riders' travel patterns," *Transportation Research Part C: Emerging Technologies*, vol. 36, pp. 1–12, 2013.
- [20] J. Y. Park, D.-J. Kim, and Y. Lim, "Use of smart card data to define public transit use in seoul, South Korea," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2063, no. 1, pp. 3–9, 2008.
- [21] S. Sun and Z. Duan, "Modeling passengers' loyalty to public transit in a two-dimensional framework: a case study in Xiamen, China," *Transportation Research Part A: Policy and Practice*, vol. 124, pp. 295–309, 2019.
- [22] S. Sun, L. Xu, Y. Yao, and Z. Duan, "Investigating the determinants to retain spurious-loyalty passengers: a data-fusion based approach," *Transportation Research Part A: Policy and Practice*, vol. 152, pp. 70–83, 2021.
- [23] S. C. Sun and D. Y. Yang, "Identifying public transit commuters based on both the smartcard data and survey data: a case study in xiamen, China," *Journal of Advanced Transportation*, vol. 2018, Article ID 9693272, 2018.
- [24] S. C. Sun, "Public transit loyalty modeling considering the effect of passengers' emotional value: a case study in xiamen, China," *Journal of Advanced Transportation*, vol. 2018, Article ID 4682591, 2018.
- [25] S. Tao, D. Rohde, and J. Corcoran, "Examining the spatial-temporal dynamics of bus passenger travel behaviour using smart card data and the flow-comap," *Journal of Transport Geography*, vol. 41, pp. 21–36, 2014.
- [26] L. Tang, Y. Zhao, K. L. Tsui, Y. He, and L. Pan, "A clustering refinement approach for revealing urban spatial structure from smart card data," *Applied Sciences*, vol. 10, no. 16, p. 5606, 2020.
- [27] Y. Long and Z. Shen, "Discovering functional zones using bus smart card data and points of interest in Beijing," in *Geospatial Analysis to Support Urban Planning in Beijing*, pp. 193–217, Springer, Berlin/Heidelberg, Germany, 2015.
- [28] C. Zhong, S. M. Arisona, X. Huang, M. Batty, and G. Schmitt, "Detecting the dynamics of urban structure through spatial network analysis," *International Journal of Geographical Information Science*, vol. 28, no. 11, pp. 2178–2199, 2014.
- [29] Y. Long and Z. Shen, *Geospatial Analysis to Support Urban Planning in Beijing*, Springer International Publishing, Berlin/Heidelberg, Germany, 2015.
- [30] Xi'an Natural Resources and Planning Bureau, *Xi'an Transportation Development Annual Report in 2018*, Xi'an Natural Resources and Planning Bureau, Xi'an, China, 2019.
- [31] P. A. P. Moran, "Notes on continuous stochastic phenomena," *Biometrika*, vol. 37, no. 1-2, pp. 17–23, 1950.
- [32] L. Anselin, "Local indicators of spatial association – LISA," *Geographical Analysis*, vol. 27, pp. 93–115, 1995.
- [33] X.-l. Ma, Y.-h. Wang, F. Chen, and J.-f. Liu, "Transit smart card data mining for passenger origin information extraction," *Journal of Zhejiang University - Science C*, vol. 13, no. 10, pp. 750–760, 2012.