

Research Article

Study of Urban System Spatial Interaction Based on Microblog Data: A Case of Huaihe River Basin, China

Yong Fan ^{1,2}, Juhui Yao,³ Zongyi He ⁴, Biao He ¹ and Minmin Li¹

¹Research Institute for Smart Cities, School of Architecture and Urban Planning, Shenzhen University, Shenzhen, China

²College of Geographic Sciences, Xinyang Normal University, Xinyang, China

³Navinfo (Xi'an) Co. Ltd., Xi'an, China

⁴School of Resource and Environmental Sciences, Wuhan University, Wuhan, China

Correspondence should be addressed to Biao He; whu_hebiao@hotmail.com

Received 6 November 2018; Accepted 17 January 2019; Published 5 February 2019

Guest Editor: Alvaro Rubio-Largo

Copyright © 2019 Yong Fan 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.

The spatial interaction of urban system has been an hot research issue in the field of urban research. In this paper, user's microblog spatial information data were used to discern the spatial structure of an urban area. Firstly, Sina Weibo microblog data for 2011–2015 were used to establish a thematic database of cities along the Huaihe River Basin, China. Secondly, network connectivity, inflow, and outflow of three indicator systems were analyzed. Finally, combining this database with socioeconomic data, experimental verification and comparative analysis were carried out. The study found that the urban spatial relation in the Huaihe River Basin has the following characteristics: the spatial difference of urban size distribution is obvious; urban layout presents a stratified aggregation phenomenon; and the high-grade cities lead the city's interaction. The research shows that this method of data mining for urban interaction in the Huaihe River Basin is valid and that this research into urban spatial patterns of river basins is applicable to other areas.

1. Introduction

In the four decades since China's reform, the urban population has increased from 170 million in 1978 to 790 million in 2016 and the urbanization rate has increased from 17.9% to 57.4% [1]. However, the level of urbanization in China is still lower than that of 70%–80% of developed countries. In the future, China will still be in the stage of rapid urbanization and development. The relationship between man and nature has become a hot topic in China at the same time. Urbanization promotes the exchange of materials, energy, personnel, and information among cities; this exchange is called urban space interaction [2]. How to effectively model and quantitatively express the spatial interaction of regional cities is a significant question and the basis of this research.

The regional urban system is a complex network [3, 4]. The connection between two nodes in this network can be represented by visible lines, such as highways and railways,

or by invisible lines, such as aviation pathways, navigation, and the Internet [5]. The flow within the system emphasizes the value of urban nodes in shaping the whole network system, which provides a theoretical framework and an important starting point for the study of urban networks [6]. Data relate to traffic, and information flow can most directly reflect the degree of urban spatial interaction [7]. Early research on this aspect studied the transmission of mail. The scope of research on transportation is more comprehensive, including the highway, railway, and aviation networks [8, 9]. In addition, the urban hierarchy can also be explored through the distribution of corporate headquarters [10] and banks [11]. The flow and consumption of network information can reflect the level of urbanization [12], which can provide basic data for geographical and sociological studies [13, 14].

The age of big data and the rapid development of LBS (location-based service) allow for urban research to be conducted by novel methods [15, 16]. With the gradual

penetration of computers into various professions, social media data [17–19] have attracted the attention of geographers due to its use in research [20]. The use of social media was one of the needs of urban communities [21]. As for the typical spatial-temporal characteristics of social media data, many data mining research papers have analyzed information, such as urban hot events [22], identifying urban Corridors [23], geography of happiness [24], whereabouts of people [25], urban function connectivity [26], urban spatial interaction [27], reevaluating urban space [28], and characterizing witness accounts [29]. Data mining [30] provides new tools, methods, and ideas for new geographical research, improving the efficiency, accuracy, and precision of the analysis [31].

Based on the data of Sina Weibo microblog [32], this paper established a thematic database of cities along the Huaihe River Basin, China. The spatial interaction of the regional urban system was analyzed from temporal and spatial characteristics of the data. Finally, combining this database with socioeconomic data, experimental verification and comparative analysis were carried out. This paper provides a reference for the exploration of the spatial relationship of regional urban systems and the driving mechanism for spatial structure of urban system.

2. Study Area

The Huaihe River Basin is located in the eastern region of China, between the Yangtze River Basin and the Yellow River Basin. It flows through the four provinces of Henan, Anhui, Jiangsu, and Shandong and covers a total area of $2.7 \times 10^5 \text{ km}^2$. The Huaihe River Basin has multiple transitional nature and society and connects coastal region and inland region, as well as being the economic integration hub of the Yangtze River and Yellow River economic belts.

The Huaihe River Basin has an enormous population of 170 million people, accounting for 12.3 percent of the total population of China. The average population density is 611 people per km^2 , which is 4.8 times of the average level across the country. Given that the Huaihe River Basin is a natural geographical unit and does not coincide with the boundaries of the administrative units, the study only selected the 26 cities whose governmental units are entirely in the Huaihe River Basin for calculation and analysis (Figure 1).

3. Data Source and Processing

3.1. Sina Weibo Microblog. The data gathered from the Sina Weibo microblog are the information released by microblog users on the social network. Compared with traditional social survey data and economic social statistics, the Sina Weibo microblog data have an unparalleled advantage in the fact that it performs well in real-time and is reliable and diverse. Currently, information can be gathered from the Sina Weibo microblog using web crawlers and the Sina Weibo microblog open platform API. This study uses the latter to obtain the relevant information for this research.

3.2. Preprocessing the Sina Weibo Microblog. First, the gathered data were cropped to the latitude and longitude range of the survey region. Each city administrative unit was divided into a 10 km-long grid, the coordinates of the grid center were obtained, and the WGS-84 geographic coordinate system was applied. Next, taking the longitude and latitude of different cities as parameters, user ID information from these cities was obtained through the nearby_timeline interface (an API interface). The user ID information includes a source (through the authorized APPKey), longitude and latitude, as well as other data. Using the user ID information as the parameter the user's historical data through the check-in interface (another API interface) were obtained. The data returned by calling the API interface were in the JSON format. In this study, a Python script was written to achieve batch data acquisition. The process of data acquisition is shown in Figure 2.

3.3. The Spatial Interaction of the Urban System. Microblog check-in data can reflect the spatial interaction of urban system; this paper adopts two methods to analyze these interactions. The first is connectivity, which is used to analyze the intensity of communication between regional cities; the second is flow, through which the user's trajectory and the interactive state among cities can be analyzed.

3.3.1. Connectivity. This method is based on the city's check-in data, from which the network connectivity among cities was analyzed, which is described as follows:

- (1) *Variable Unitization.* After variable unitization, a matching matrix was established regarding the city check-in data using the following equation:

$$V_{ij}0 = \frac{V_{ij}1}{\sum_{j=1}^{25} V_{ij}}, \quad (1)$$

where $V_{ij}0$ is a microblog user in city i who signed into city j after unitization; $V_{ij}1$ represents a microblog user of i city who signed in city j after data acquisition; and $\sum V_{ij}$ represents the sum of the check-in data of user from city i in city j .

- (2) *Urban External Connection Index.* The urban external connection index represents the difference between the total number of check-ins in other cities and those in city i . This is calculated using the following equation:

$$N_i0 = \sum_{j=1}^{25} V_{ij}0 - V_{ii}0, \quad (2)$$

where N_i0 is the external connection index of city i , which reflects the check-in data (V) of city i in the other 25 cities; $V_{ii}0$ represents the check-in data of city i after unitization; and $V_{ij}0$ is the sum of check-in data of microblog users from city i who signed in other cities after unitization.

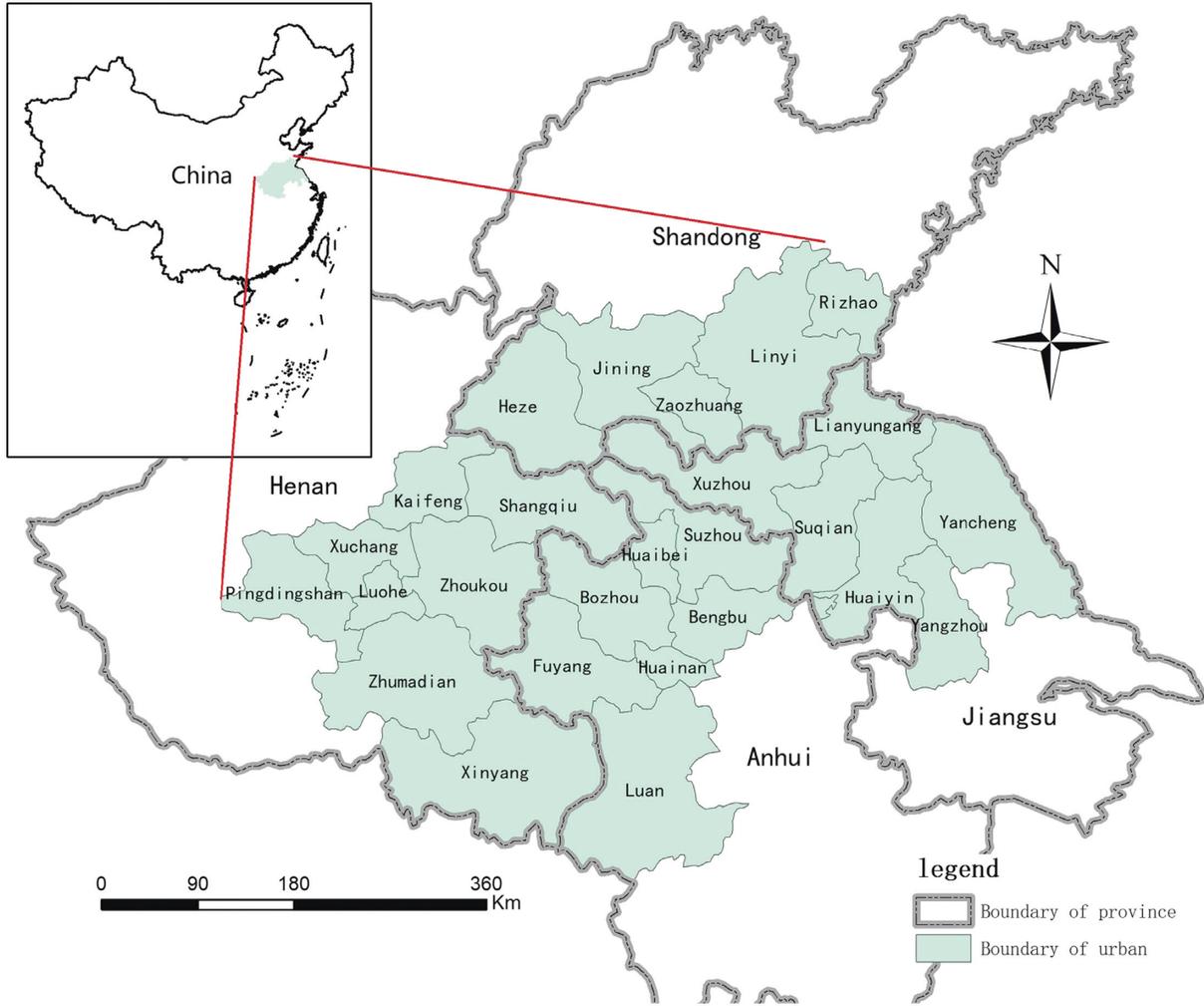


FIGURE 1: Study area.

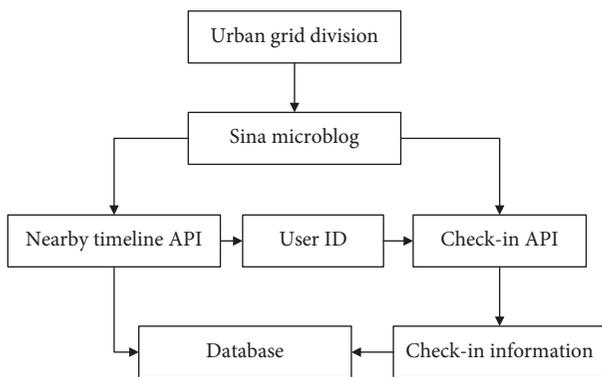


FIGURE 2: The process of sina Weibo microblog data acquisition.

If $N_{i0} > 0$, users of city i sign into the microblog more when not in city i ; if $N_{i0} < 0$, users registered to city i check into Sina Weibo locally. $N_{i0} = 0$ means that users from city i are equally likely to sign into the microblog whether they are within city i or away from the city.

(3) *Urban Network Connectivity*. First, the standardized check-in data of city i in city j and the standardized check-in data of city j in city i are multiplied together to obtain the urban network connectivity. Then, these data are standardized to produce the standardized network connectivity. This is summarized in the following equation:

$$R_{ij} = V_{ij0} * V_{ji0},$$

$$R_{ij0} = R_{ij} * \left(\frac{100}{\text{Max}(R_{ij})} \right), \tag{3}$$

where R_{ij0} is the standardized network connectivity between city i and city j ; R_{ij} is the network connectivity between city i and city j ; V_{ij0} is the standardized value of friends of city j residents who live in city i ; $\text{Max}(R_{ij})$ is the maximum network connectivity value; and R_{ij0} reflects interconnectedness of each city's information.

(4) *Network Connectivity in Each City*. M_{i0} is the standard network connectivity of city i with itself.

M_i0 is the difference between the standard network connectivity of city i and city j and the standard network connectivity of city i itself. This is calculated using the following equation:

$$M_i0 = \sum_{j=1}^{25} R_{ij}0 - R_{ii}0, \quad (4)$$

where M_i0 reflects the linkage strength of city i within the network system.

3.3.2. Inflow Rating and Outflow Rating. The inflow and outflow ratings reflect the attractiveness and interaction of cities, respectively. The inflow rating is the total number of other cities signed into the city. The greater the inflow, the higher the attractiveness and degree of interaction the city has. The outflow rating reflects the total number of users registered to the city signed into other cities.

The greater the outflow rating, the greater the output of the city, which indicates the city has a higher degree of interaction with other cities.

Formula (5) is used to calculate the inflow rating of city i , where V_{i1} indicates the number of check-ins from users registered to city j in city i :

$$V_{i1} = \sum_{j=1}^{25} V_{ij}. \quad (5)$$

Formula (6) is used to calculate the outflow rating of city i , where V_{i2} indicates the number of check-ins from users registered to city j in city i :

$$V_{i2} = \sum_{j=1}^{25} V'_{ij}. \quad (6)$$

To further analyze the reasons for flow among cities, the ratio of inflow to outflow is introduced, which can reflect the population flow among cities. This is calculated using the following equation:

$$R = \frac{V_{i1}}{V_{i2}}. \quad (7)$$

4. Results

4.1. Connectivity Analysis. Based on the calculation of urban connectivity, the microblog check-in data in the study area can be calculated. From this, the urban external connectivity index and network connectivity are obtained. This data are summarized in Table 1.

In Table 1, the N_i0 value of Luohe and Kaifeng is 1, which shows that there are more check-ins from outside users into both Luohe and Kaifeng. The N_i0 value of Rizhao is -0.60 , which is the lowest in the Huaihe River Basin. This indicates that check-in data of Rizhao consist of nearly only local users. In addition, the M_i0 value of Bengbu is -1543.59 , which is the minimum in the Huaihe River Basin. Bengbu's external relations with the basin as a whole are relatively weak. The M_i0 of Luohe is 291.15, which is the largest in the Huaihe River

TABLE 1: External connection index and network connectivity.

Name	N_i0	M_i0
Luohe	1.00	291.15
Kaifeng	1.00	203.67
Huainan	0.97	159.24
Bozhou	0.95	120.38
Yangzhou	0.86	-84.64
Fuyang	0.62	-47.69
Xinyang	0.42	-83.31
Lianyungang	0.39	-75.32
Jining	0.32	-74.07
Suzhou	0.23	-86.43
Huaiyin	-0.01	-96.93
Zhumadian	-0.08	-97.96
Suqian	-0.15	-96.90
Heze	-0.16	-98.96
Zhoukou	-0.22	-98.30
Bengbu	-0.25	-1543.59
Xuchang	-0.30	-96.77
Liuan	-0.31	-98.11
Zaozhuang	-0.31	-98.63
Pingdingshan	-0.36	-97.27
Huaibei	-0.50	-95.69
Shangqiu	-0.51	-97.60
Linyi	-0.53	-97.88
Xuzhou	-0.53	-96.98
Yancheng	-0.59	-90.80
Rizhao	-0.60	-99.02

Basin. It shows that Luohe has a strong external relationship with the whole basin. Luohe is located in the middle of the Henan Province, next to the provincial capital Zhengzhou. The city has convenient transportation, prosperous economy, well-developed tourism industry, and large population mobility; these factors allow for more users to travel to other locations and increase the M_i0 value of Luohe.

By analyzing the relationship between the external and internal connectivity of a city (Figure 3), the following three conclusions were reached:

- (1) The internal urban network connectivity is consistent with the urban external connectivity: the lower the city ranks in terms of internal connectivity, the lower the city's level of external connections with the rest of the network. Among the 26 cities, Luohe, Kaifeng, Huainan, and Bozhou comprised the top four. The positive values for these cities indicate that the intensity of external connections between the four cities is greater than that within the cities. The remaining 22 cities have negative values, which show that these cities have low external connections and that communication within these cities is greater than that with the external network. The network connectivity of Fuyang and Bengbu does not correspond to the urban city hierarchy which shows that, although the interconnectedness of these 2 cities with the other cities in the basin is not high, their absolute degree of network connectivity is relatively high. This is mainly due to the strong network connectivity in a small area. This phenomenon is most obvious in Bengbu.

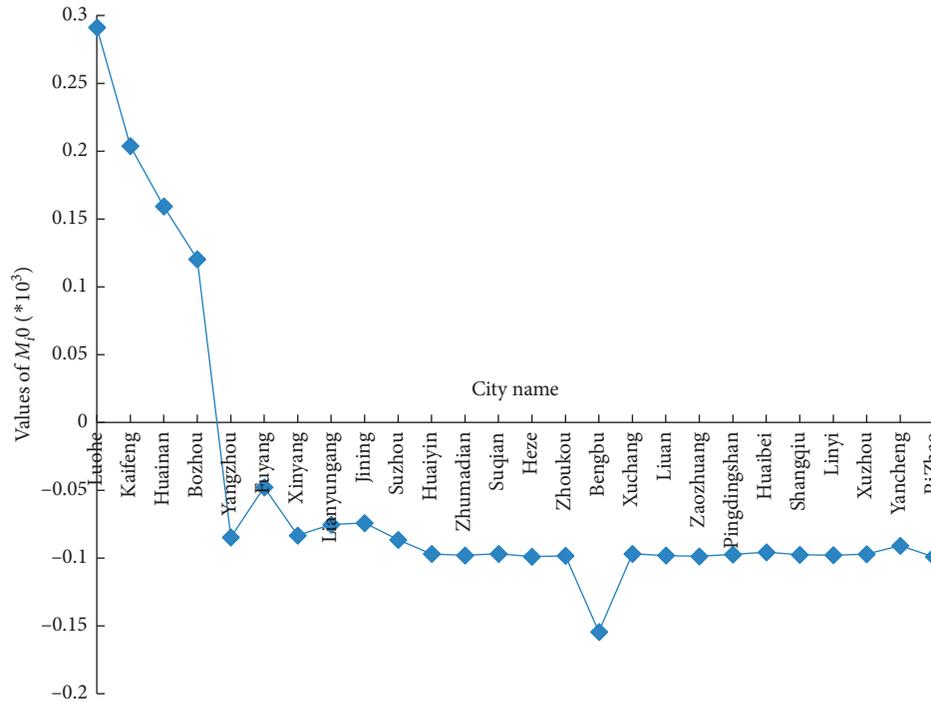


FIGURE 3: Relationship between urban network connectivity and urban hierarchy.

- (2) Foreign contact intensity is related to a city’s economic level: the standardized network connectivity (M_i0) reflects the intensity of a city’s linkage to the entire basin system. M_i0 is derived from the difference between the standardized network connectivity value of city i with the other cities and the value of the city i within itself. M_i0 reflects the strength of the external connection of city i . Based on the classification of natural breakpoints (Jenks classification), the 26 cities are divided into 5 grades, which are shown in Table 2. According to the statistical results of the four provinces GDP (Jiangsu, Shandong, Anhui, and Henan), the 26 cities are ranked in the order of Jiangsu, Shandong, Henan, and Anhui. The five cities of Luohe, Kaifeng, Huainan, Bozhou, and Fuyang in Henan and Anhui were ranked the highest. This is because the population of the two cities generally interacts on a local scale.
- (3) Significant difference in connection intensity between cities: the network connectivity of 26 cities was sorted were obtained. From the data, it was found that the strength of the city-to-city relationship varies greatly. There are 225 sets of data between 0 and 1, occupying 69.23% of the total, and 62 sets of data between 1 and 10, occupying 19.08% of the total. Moreover, 31 sets of data between 10 and 100 accounted for 9.54% of the data, while 16 sets of 100–1000 accounting for 4.92%.

4.2. Urban Interaction Analysis. According to the 26 cities’ check-in data and the user’s mobile track, the urban interaction of the 26 cities can be analyzed.

TABLE 2: Classification of foreign contact intensity.

Grade	City	Connectedness value
1	Luohe, Kaifeng, Huainan, Bozhou	-47.69 to 291.45
2	Fuyang	-74.07 to -47.69
3	Jining, Lianyungang, Xinyang, Yangzhou, Suzhou	-90.80 to -74.07
4	Rizhao, Heze, Zaozhuang, Zhoukou, Liu'an, Zhumadian, Linyi, Shangqiu, Pingdingshan, Xuzhou, Huaiyin, Suqian Xuchang, Huaibei, Yancheng, Suzhou, Yangzhou, Xinyang	-154.36 to -90.80
5	Bengbu	-154.36

4.2.1. Classification of Inflow Ratings. The inflow rating of the city reflects the total number of microblog entries from other cities signed in the city. Taking city i as the research object, the check-in data of city j in city i are collected, and then the total inflow of city i is obtained. According to the natural breakpoint, the 26 cities are divided into 5 grades, as shown in Table 3.

From Table 3, we can see that the inflow ratings of the cities have the following characteristics:

- (1) The difference in inflow level is large: the largest number of check-ins to a city is 96,000, while the smallest city had less than 10,000 check-ins.
- (2) The fourth city rank had the most of cities; most cities had between 10,000 and 20,000 check-ins.
- (3) The check-in data are unevenly distributed: The data are mainly distributed in the northeast, southeast, and east of the basin, mainly in the Jiangsu and Shandong provinces. The amount of data in the Anhui and Henan provinces is relatively small.

TABLE 3: Classification of inflow rating.

Grade	City	Inflow rating
1	Jining, Xuzhou, Bengbu, Yancheng	60,001–96,000
2	Shangqiu, Huaibei, Huaiyin	40,001–60,000
3	Kaifeng, Linyi, Lianyungang, Yangzhou, Liuan	20,001–40,000
4	Pingdingshan, Xuchang, Luohe, Xinyang, Zhoukou, Fuyang, Bozhou, Huainan, Suzhou, Zaozhuang, Suqian, Rizhao	10,001–20,000
5	Heze, Zhumadian	10,000

4.2.2. *Classification of Outflow Ratings.* The outflow rating reflects the total number of users from a city that signed into other cities. Similarly to the inflow rating, the data were classified according to the natural breakpoints, and the results shown in Table 4 were obtained.

From Table 4, it can be seen that the outflow has the following characteristics.

- (1) The difference in outflow level varies significantly across the basin: the largest number of check-ins from one city is more than 100,000, while the smallest city has less than 10,000.
- (2) The fourth city rank is the largest in terms of outflow: most cities have between 10,000 and 30,000 check-ins in other cities.
- (3) The spatial distribution of the outflow data is uneven: the data are mainly distributed in the central and western regions of the basin. Out of all the provinces, the outflow from the Jiangsu province is the largest, Shandong province is unevenly distributed, and Henan and Anhui provinces have more homogeneous outflow spatial distributions.

4.2.3. *Ratio of Inflow and Outflow.* The ratio of inflow and outflow can directly reflect the flow direction of population between cities and can also help to further analyze the driving factors of population flow between cities. The ratio of inflow and outflow can reflect the attractiveness of a city. The higher the ratio, the greater the attractiveness of the recipient city, or less attractive the city from which the people are migrating. According to natural breakpoint classification, the 26 cities are divided into 5 grades (Table 5).

From Table 5, the following ratio of inflow and outflow characteristics can be determined:

- (1) The difference between the ratios of inflow and outflow between cities is large: the largest ratio is 1244 times more than the smallest.
- (2) Most of the cities had inflow/outflow ratios below 1, and these cities were distributed more evenly. This shows that the inflow of most cities is less than that of the outflow.
- (3) The ratio of inflow and outflow in Luohe, Kaifeng, and Huainan are relatively large, which indicates that the inflow and outflow data of these three cities are significantly different from those of the other cities in the river basin.

TABLE 4: Classification of outflow rating.

Grade	City	Outflow rating
1	Yangzhou	80,001–102,087
2	Xinyang, Bozhou, Jining, Lianyungang	50,001–80,000
3	Luohe, Kaifeng, Fuyang, Linyi, Huaiyin Pingdingshan, Xuchang, Zhoukou,	30,001–50,000
4	Zhumadian, Shangqiu, Heze, Suzhou, Xuzhou, Zaozhuang, Suqian, Huainan, Bengbu, Luan, Yancheng	10,001–30,000
5	Huaibei, Rizhao	10,000

TABLE 5: Ratio of inflow and outflow.

Grade	City	Inflow and outflow ratio
1	Luohe	>300
2	Kaifeng, Huainan	51–300
3	Bozhou, Yangzhou Xinyang, Fuyang, Suzhou, Jining,	6–50
4	Lianyungang Pingdingshan, Xuchang, Zhoukou, Zhumadian, Shangqiu Heze, Huaibei, Bengbu, Lu'an, Huaiyin,	1–5
5	Suqian, Xuzhou, Zaozhuang, Linyi, Rizhao, Yancheng	<1

5. Discussion

In this paper, we used microblog data to effectively model and quantitatively express the spatial interaction of regional cities. Microblog data can not only reflect mobile information of users' tracks but also reflect users' travel and activity rules. Does socioeconomic status affect users' travel and activity rules? How does it affect? In order to answer these questions, this paper analyzes the correlation among urban microblog registration data, urban population, and GDP with the statistical data.

5.1. *The Relationship between Urban Interaction and Social Economy.* Take the GDP and the urban population as the abscissa, respectively, and the total city check-in data as the ordinates, Figures 4 and 5 are obtained.

From Figures 4 and 5, it can be concluded that there is no positive correlation among the total amount of urban check-in, the economic links in the entity and the total population. To a certain extent, this reflects the degree of urban interaction is a mechanism of the joint action of the economic level and population factors of the city. The city's largest inflow and local check-in data are the side portraits of the city's attractiveness index, which determines the level of city interaction. Cities with high interactive levels, such as Xuzhou and Yancheng, have high inflows and local check-in volumes, and their GDP is also higher in the provinces. Suqian, Suzhou, Zhoukou, Pingdingshan, Heze, Zhumadian, and other cities are based on the local sign of data, so their inflow is less and the interaction between cities is not strong. From the data of the Statistical Yearbook of 2015, it is not difficult to find that the cities with strong interactive ranks have higher GDP levels. Therefore,

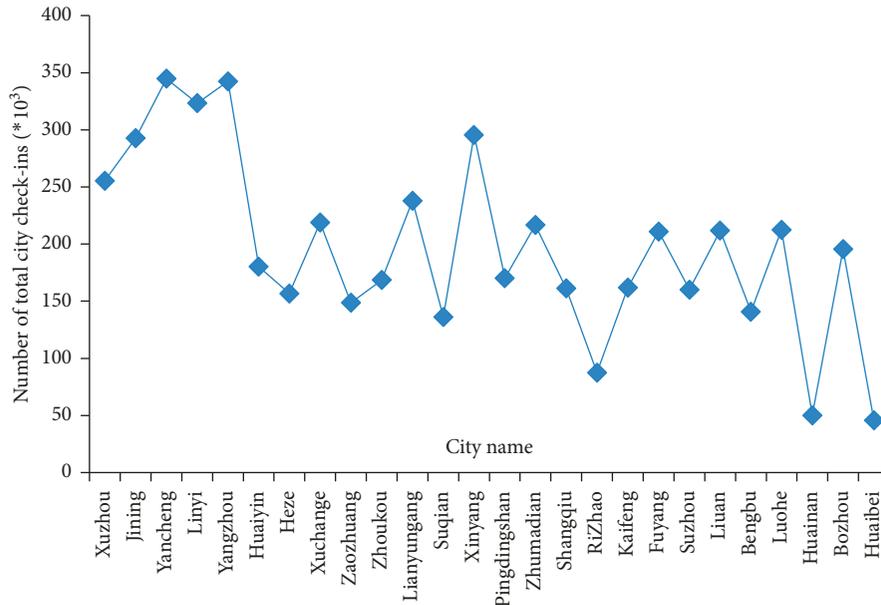


FIGURE 4: Total attendance ordered by GDP.

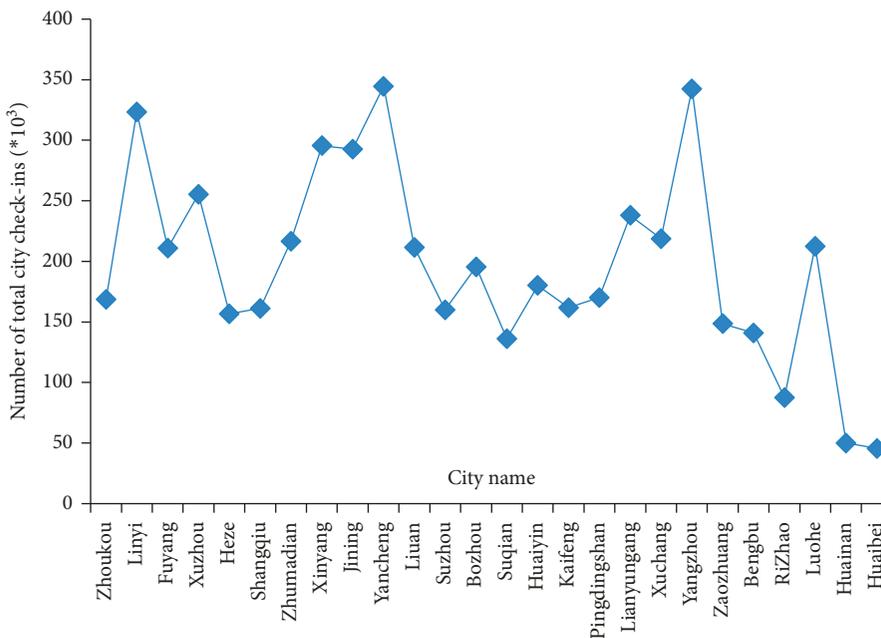


FIGURE 5: Total attendance ordered by population.

if we want to improve the city’s GDP, we must strengthen the interaction between cities.

5.2. *Microblog Data and Traditional Data.* As a class of big data, microblog data contain user’s location information which is different from traditional data, such as highway data and railway data. The spatial structure of urban system is the traditional field of human geography research. With the development of new urbanization and coordinated development of urban and rural areas, the research of urban system structure is developed from the morphological

structure to social structure, cultural structure, flow structure, functional structure, and other fields [33]. Therefore, we need new perspectives and methods to support the study of the spatial structure of urban system.

The key to analyze the characteristics of user movement trajectories in specific areas is data acquisition. Traditional methods of group analysis in specific regions have three steps. First, define the group in the region; second, conduct a sample survey of the people in the region by using a questionnaire; third, analyze the statistical data. Traditional sampling survey could not get accurate facts to express data and always consumes large amount of time, manpower, and

material resources. Microblog data are the information released by microblog users in social networks. The content of microblog data has a strong real-time reliability and diversity, which has the advantage that traditional data do not have. In the case of scientific research, it is more accurate and persuasive to further analyze the user's information without revealing the user's personal information and privacy. Microblog data and traditional data have their own advantages. If we combine the two factors to analyze user behavior, the result of analysis will be more reasonable. The paper adopts this method when analyzing the correlation between microblog data and socioeconomic statistics.

The quality of data directly determines the accuracy of data analysis results. The data acquired by microblog should follow certain rules and should show its three characteristics of continuity, integrity, and validity. Therefore, when accessing data, we should pay attention to the time and number of interface calls, as well as the integrity and repeatability checking of data. The check-in information obtained in the study includes check-in city code, provincial code, check-in longitude and latitude coordinates, check-in date, and check-in time. Get microblog data and establish a database corresponding to the data field. After entering the warehouse, the missing data and the noise data are removed to ensure the completeness and uniformity of the data.

5.3. River Basin Perspective Study of Urban Spatial Interaction.

In recent years, reunderstanding of the relationship between man and nature has become a hot topic in various fields. The city is an important habitat of human activities. The spatial and temporal structure changes of urban system imply the dynamic interaction between human and nature. Studying the spatial interaction of regional urban system and revealing the driving forces in the process of its change will help us to understand the temporal and spatial evolution of the urban system scientifically. The above research can provide the basis of space research for the sustainable development of human society and also provide a technical route reference for the spatial layout and optimization of the cities, which has a certain theoretical and practical significance.

As a complete physical and geographical unit, the River Basin crosses the administrative boundary and could reflect the regional human-land relationship more naturally and truly. The study of spatial interaction between cities by the river basin excludes the human subjective constraints, which more objectively reflect the spatial relationship characteristics of the regional urban system. Therefore, the river basin is an ideal area to explore the evolution rules of the relationship between mankind and land. The Huaihe River Basin, as the transition zone of China's urban system, has dual transitional nature and social elements [34]. Therefore, the research takes the Huaihe River Basin as a case area to study the spatial relationship of China's regional urban system. Its perspective is unique.

6. Conclusions

This paper obtains the user microblog information through the sina microblog open platform and studies the urban spatial pattern and urban interaction by means of statistical analysis and spatial analysis. This paper takes the Huaihe basin as the case area to verify it. The main research conclusion as follows:

- (1) The data interface provided by microblog platform can study the urban spatial pattern. The user trajectory of microblog data can explore the spatial relationship of regional cities, and data acquisition and data quality evaluation can meet the research requirements.
- (2) Based on microblog data, the spatial and temporal characteristics of urban system spatial pattern in Huaihe River Basin are analyzed from network connectivity and urban interaction. The study found that the urban spatial relation in the Huaihe River Basin has the following characteristics: the spatial difference of urban size distribution is obvious; urban layout presents a stratified aggregation phenomenon; and the high-grade cities lead the city's interaction.

As for the application of microblog data in urban research, the current mainly focus on information text, social relations, and other aspects. The research is mainly about event detection and hot spot exploration. The combination of big data thinking and data mining technology will have more research findings in the study of urban problems.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

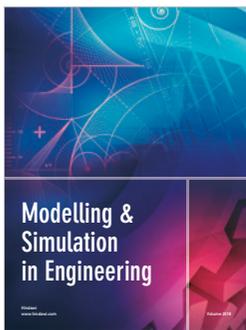
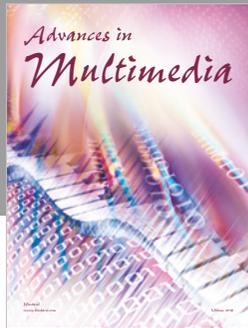
Acknowledgments

This work was supported by the Natural Science Foundation of China (Grant no. 41701187) and project funded by China Postdoctoral Science Foundation (Grant nos. 2018M640813 and 2018M633108).

References

- [1] Chinese National Bureau of Statistics, 2017, in Chinese, <http://www.stats.gov.cn>.
- [2] Y. Liu, L. Gong, and Q. Tong, "Quantifying the distance effect in spatial interactions," *Acta Scientiarum Naturalium Universitatis Pekinensis*, vol. 50, no. 3, pp. 526–534, 2014.
- [3] Y. Chen, "Zipf's law, hierarchical structure, and shuffling-cards model for urban development," *Discrete Dynamics in Nature and Society*, vol. 2012, Article ID 480196, 21 pages, 2012.

- [4] H. Zhang, L. Zhu, and S. Xu, "Modeling the city distribution system reliability with bayesian networks to identify influence factors," *Scientific Programming*, vol. 2016, Article ID 7109235, 2016.
- [5] F. Zhen, B. Wang, and Y.-x. Chen, "China's city network characteristics based on social network space: an empirical analysis of sina microblog," *Acta Geographica Sinica*, vol. 67, no. 8, pp. 1031–1043, 2012.
- [6] F. Giorgio and S. Gianluca, "Human-mobility networks, country income, and labor productivity," *Network Science*, vol. 3, no. 3, pp. 377–407, 2015.
- [7] D. Dandan, Z. Chunshan, and Y. Changdong, "Spatial-temporal characteristics and factors influencing commuting activities of middle-class residents in Guangzhou city, China," *Chinese Geographical Science*, vol. 26, no. 3, pp. 410–428, 2016.
- [8] W. Wu, Y.-h. Cao, S.-b. Liang, and W.-d. Cao, "The accessibility pattern of railway passenger transport network in China," *Geographical Research*, vol. 28, no. 5, pp. 1389–1400, 2009, in Chinese.
- [9] J.-f. Xue, "Hierarchical structure and distribution pattern of Chinese urban system based on aviation network," *Acta Geographica Sinica*, vol. 27, no. 1, pp. 23–32, 2008, in Chinese.
- [10] Z.-f. Jin, "On structural properties of transnational urban network based on multinational enterprises network in China: as the case of link with South Korea," *Acta Geographica Sinica*, vol. 29, no. 9, pp. 1670–1682, 2010, in Chinese.
- [11] J. Yin, F. Zhen, and C.-h. Wang, "China's city network pattern: an empirical analysis based on financial enterprises layout," *Economic Geography*, vol. 31, no. 5, pp. 754–759, 2011, in Chinese.
- [12] Z.-w. Sun, J.-l. He, and F.-j. Jun, "The accessibility and hierarchy of network cities in the global Internet," *Economic Geography (Chinese)*, vol. 30, no. 9, pp. 1449–1455, 2010.
- [13] M. Stephens and A. Poorthuis, "Follow thy neighbor: connecting the social and the spatial networks on Twitter," *Computers Environment and Urban Systems*, vol. 53, no. 3, pp. 331–346, 2014.
- [14] Y. Takhteyev, A. Gruzd, and B. Wellman, "Geography of twitter networks," *Social Networks*, vol. 34, no. 1, pp. 73–81, 2012.
- [15] Z. Liu, Y. Jia, and X. Zhu, "Deployment strategy for car-sharing depots by clustering urban traffic big data based on affinity propagation," *Scientific Programming*, vol. 2018, Article ID 3907513, 9 pages, 2018.
- [16] M. M. Rathore, A. Ahmad, A. Paul, and S. Rho, "Urban planning and building smart cities based on the Internet of Things using Big Data analytics," *Computer Networks*, vol. 101, pp. 63–80, 2016.
- [17] D. Arribas-Bel, K. Kourtit, P. Nijkamp et al., "Cyber cities: social media as a tool for understanding cities," *Applied Spatial Analysis and Policy*, vol. 8, no. 3, pp. 231–247, 2015.
- [18] J. Cinnamon and N. Schuurman, "Confronting the data-divide in a time of spatial turns and volunteered geographic information," *GeoJournal*, vol. 78, no. 4, pp. 657–674, 2012.
- [19] A. Stefanidis, A. Crooks, and J. Radzikowski, "Harvesting ambient geospatial information from social media feeds," *GeoJournal*, vol. 78, no. 2, pp. 319–338, 2011.
- [20] S. Elwood, "Geographic information science: emerging research on the societal implications of the geospatial web," *Progress in Human Geography*, vol. 34, no. 3, pp. 349–357, 2009.
- [21] N. Ratnasari, E. D. Candra, D. H. Saputra, and A. P. Perdana, "Urban spatial pattern and interaction based on analysis of nighttime remote sensing data and geo-social media information," *IOP Conference Series: Earth and Environmental Science*, vol. 47, article 012038, 2016.
- [22] A. Schulz, E. L. Mencia, and B. Schmidt, "A rapid-prototyping framework for extracting small-scale incident-related information in microblogs: application of multi-label classification on tweets," *Information Systems*, vol. 57, pp. 88–110, 2015.
- [23] Z. Jiang, M. Evans, D. Oliver et al., "Identifying K Primary Corridors from urban bicycle GPS trajectories on a road network," *Information Systems*, vol. 57, pp. 142–159, 2016.
- [24] L. Mitchell, M. R. Frank, K. D. Harris, P. S. Dodds, and C. M. Danforth, "The geography of happiness: connecting twitter sentiment and expression, demographics, and objective characteristics of place," *PLoS One*, vol. 8, no. 5, Article ID e64417, 2013.
- [25] E. Steiger, R. Westerholt, B. Resch, and A. Zipf, "Twitter as an indicator for whereabouts of people? Correlating Twitter with UK census data," *Computers, Environment and Urban Systems*, vol. 54, pp. 255–265, 2015.
- [26] Y. Shen and K. Karimi, "Urban function connectivity: characterisation of functional urban streets with social media check-in data," *Cities*, vol. 55, pp. 9–21, 2016.
- [27] Z.-wei Sui, L. Wu, and Y. Liu, "Study on interactive network among Chinese cities based on the check-in dataset," *Geography and Geo-Information Science (Chinese)*, vol. 29, no. 6, pp. 1–5, 2013.
- [28] G. Wessel, C. Ziemkiewicz, and E. Sauda, "Reevaluating urban space through tweets: an analysis of Twitter-based mobile food vendors and online communication," *New Media and Society*, vol. 18, no. 8, pp. 1636–1656, 2016.
- [29] M. Truelove, M. Vasardani, and S. Winter, "Towards credibility of micro-blogs: characterising witness accounts," *GeoJournal*, vol. 80, no. 3, pp. 1–21, 2015.
- [30] J. Capdevila, M. Arias, and A. Arratia, "GeoSRS: a hybrid social recommender system for geolocated data," *Information Systems*, vol. 57, pp. 111–128, 2016.
- [31] D. Kotzias, T. Lappas, and D. Gunopulos, "Home is where your friends are: utilizing the social graph to locate twitter users in a city," *Information Systems*, vol. 57, pp. 77–87, 2016.
- [32] Chinese Weibo, 2018, in Chinese, <https://weibo.com>.
- [33] Z. Wang, X. Ye, J. Lee, X. Chang, H. Liu, and Q. Li, "A spatial econometric modeling of online social interactions using microblogs," *Computers, Environment and Urban Systems*, vol. 70, pp. 53–58, 2018.
- [34] Y. Fan, G. Yu, and Z. He, "Origin, spatial pattern, and evolution of urban system: testing a hypothesis of 'urban tree'," *Habitat International*, vol. 59, pp. 60–70, 2017.




Hindawi

Submit your manuscripts at
www.hindawi.com

