





Article

# Social Sensing of the Imbalance of Urban and Regional Development in China Through the Population Migration Network around Spring Festival

Ruoxin Zhu <sup>1,\*</sup>, Diao Lin <sup>1</sup>, Yujing Wang <sup>2,3</sup>, Michael Jendryke <sup>2,4</sup>, Rui Xin <sup>2</sup>, Jian Yang <sup>5</sup>, Jianzhong Guo <sup>6</sup> and Liqiu Meng <sup>1</sup>

<sup>1</sup> Chair of Cartography, Technical University of Munich, 80333 Munich, Germany; diao.lin@tum.de (D.L.); liqiu.meng@tum.de (L.M.)

<sup>2</sup> School of Resource and Environmental Sciences, Wuhan University, Wuhan 430079, China; afreeboat@whu.edu.cn (Y.W.); mjendryke@whu.edu.cn (M.J.); xinrui@whu.edu.cn (R.X.)

<sup>3</sup> Key laboratory of Geographic Information Systems, Ministry of Education, Wuhan University, Wuhan 430079, China

<sup>4</sup> State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

<sup>5</sup> School of Geospatial Information, Information Engineering University, Zhengzhou 450052, China; jian.yang@tum.de

<sup>6</sup> College of Environment and Planning, Henan University, Kaifeng 475004, China; guo\_jianzhong420@sohu.com

\* Correspondence: ruoxin.zhu@tum.de; Tel.: +49-152-5723-5586

Received: 1 April 2020; Accepted: 22 April 2020; Published: 23 April 2020



**Abstract:** Regional development differences are a universal problem in the economic development process of countries around the world. In recent decades, China has experienced rapid urban development since the implementation of the reform and opening-up policy. However, development differs across regions, triggering the migration of laborers from underdeveloped areas to developed areas. The interaction between regional development differences and Spring Festival has formed the world's largest cyclical migration phenomenon, Spring Festival travel. Studying the migration pattern from public spatiotemporal behavior can contribute to understanding the differences in regional development. This paper proposes a geospatial network analytical framework to quantitatively characterize the imbalance of urban/regional development based on Spring Festival travel from the perspectives of complex network science and geospatial science. Firstly, the urban development difference is explored based on the intercity population flow difference ratio, PageRank algorithm, and attractiveness index. Secondly, the community detection method and rich-club coefficient are applied to further observe the spatial interactions between cities. Finally, the regional importance index and attractiveness index are used to reveal the regional development imbalance. The methods and findings can be used for urban planning, poverty alleviation, and population studies.

**Keywords:** crowdsourced geographical data; Spring Festival travel; network analysis; community detection; rich-club phenomenon; regional development differences

## 1. Introduction

Regional development differences have always been one of the core issues in urban study and a universal problem in the economic development process of countries around the world [1–3]. The unbalanced development theory reveals the unbalanced development of economic sectors or

industries and emphasizes the optimization of resource allocation [4]. The circular and cumulative causation theory indicates that the economic development process does not occur simultaneously and spread uniformly in space, but starts from some areas with better conditions [5]. Based on the viewpoints of the center-periphery theory and the growth pole theory, in any country's regional system, a few regions with good location conditions and good industries will grow into economic growth poles through the polarization effect [6,7]. Then, through the back-wash effect and the spread effect of the growth pole, they affect and drive the development of surrounding areas and other industries. The hypothesis of the inverted-U-shaped relationship between spatial inequality and economic development shows that in the initial stage of development, the back-wash effect will be more obvious, and a large amount of capital, technology, talent, and other production factors will gather at the growth poles [8]. Therefore, there must be a small number of cities that have gradually developed into regional hubs, attracting migrants from the surrounding areas.

From the unbalanced development theory, it can be inferred that population migration has a strong internal relationship with regional development differences, and laborers seeking higher living standards will move from areas with relatively poor development to areas with better economic conditions. Currently, many scholars measure the level of urban development by constructing indicators based mainly on the city's ecology, population, economy, education, etc., and less on the interaction between cities. However, spatial interactions between cities, such as transportation networks, trade, and population movements, can also serve as evidence of urban development differences [9]. In particular, population migration, as a social expression of the spatial interaction between the source city and the destination city, can intuitively reflect the regional differences in urban development. Within a country, relatively developed cities can provide higher incomes and better living standards and thereby attract laborers from underdeveloped regions. Therefore, studying the migration pattern from the spatiotemporal behavior of the public can contribute to understanding the differences in urban/regional development.

At present, the relevant regional unbalanced development theories have been used to interpret population migration. However, few studies have revealed the unbalanced regional development from population migration patterns. The link between population migration patterns and urban/regional development needs to be explored with appropriate theories and data. Although population migration exists in physical space, it can be mapped to a geospatial network, with cities as network nodes and the direction and volume of population flow as the directions and weights of the network edges. This paper constructs a geospatial migration network based on the population migration flow and explores the underlying structural and functional characteristics. The PageRank algorithm and optimized attractiveness index are used to evaluate urban/regional importance and attraction, respectively. The community detection method and the rich-club coefficient are further used to observe the spatial interaction pattern between cities. The proposed method is applied to Chinese Spring Festival migration for empirical research, and analysis results are compared with urban economic statistics and government development plans to reveal the relationship between the migration patterns and urban/regional development. Methodologically, this study constructs an analytical framework devoted to unveiling the relationship between the spatial configuration of the population migration network (PMN) and urban/regional development from multiple geospatial network aspects, including urban development differences, urban spatial interactions, and regional development imbalance. This universal methodological framework can be applied to other related fields (e.g., social networks, business trade, and traffic networks) to identify the differences and connections between nodes and subgroups in a network.

Concerning the organization of this paper, we describe the study data and verify the rationality of spatial sampling in Section 2. Section 3 introduces the methodological framework for exploring urban/regional development differences based on the PMN around Spring Festival and describes the corresponding methods in detail. The research results are presented in Section 4 and analyzed in

Section 5 in conjunction with relevant research findings and related development planning policies in China. Finally, Section 6 concludes the paper and proposes potential directions for further exploration.

## 2. Chinese Spring Festival Travel

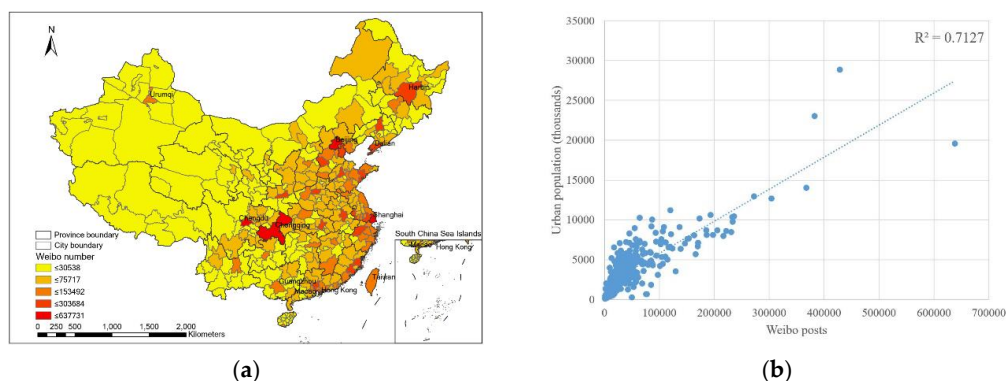
China is a multiethnic country with a vast territory, large population, and diverse climate and natural conditions, in which 18.59% of the world's population lives on 6.3% of the world landmass (Worldometer's data). In recent decades, China has experienced rapid urban development as it implements a series of reform and opening-up policies [10]. However, due to the influence of the historical-geographical environment, natural resource conditions, transportation, government policies, etc., urban development differs significantly across the regions of China. The imbalance in urban development leads to various social and economic problems that cannot be ignored, such as the influx of migrant workers in developed cities, urban contraction and population loss in underdeveloped regions, and challenges in providing education for left-behind children [11]. In China, large-scale intercity population flow usually occurs around holidays and is dominated by tourism. However, the most significant migration phenomenon in China, Spring Festival travel (SFT), is different. The Spring Festival is a traditional Chinese festival and is regarded as an occasion for family reunion. Therefore, many migrants leave the city where they work and go back to their hometowns for the family reunion before the Spring Festival, and after that, they return to the cities where they live. The root cause of this phenomenon of SFT is the uneven development of the regional economy and deep-rooted cultural tradition. The migration pattern uncovered by the population flow of the SFT rush season can be used as evidence of uneven regional development in China.

Research on Chinese population migration is ongoing. Many scholars have discussed the spatial autocorrelation effect and network autocorrelation effect in interprovincial migration flows based on the fifth and sixth census data obtained in 2000 and 2010, respectively [12]. However, census data are counted every ten years, and other multisource data are needed as a time-sensitive supplement to support ongoing research on population migration. In recent years, crowd-sourced geographical data have been generated as a result of speedy scientific-technical progress, especially in mobile positioning technology, internet technology, and communication technology, triggering a data-driven channel for sensing human spatial behavior [13]. Crowd-sourced geographical data (e.g., mobile check-in data, cellular signaling data, and taxi trajectory data) are rich in information, low cost, and abundant [14]. Such data have been widely used in sensing the geographical environment [15], recognizing urban structure and functional areas [16], planning urban development [17], assisting sustainable economic development [18], perceiving geographical events [19,20], and crowdmapping [21]. Therefore, social sensing based on crowd-sourced geographical data provides a practical approach to explore the spatial behavior of the public and reveal geographical features of the socioeconomy [22].

In addition to the traditional decennial census data, limited population flow data have been made publicly available by several location-based services (LBS) platforms, allowing a data-driven social sensing approach to observe population migration around Spring Festival. Three mainstream LBS platforms (i.e., Baidu, Tencent, and Qihoo) were compared in reference [23]. Qihoo does not provide city-level population flow data, while Baidu and Tencent provide only the top 10 inflows and top 10 outflows for each city. For a country such as China with hundreds of cities, such data cannot allow the full exploration of the pattern of population migration across cities. By contrast, geotagged social media data provide another way to perceive human behavior. As information sharing (e.g., personal life updates and information sharing) and social connection are two primary motivations for using social media services [24], the spatial-temporal information embedded in social media data provides valuable indicators for sensing human movements. Sina Weibo is the most popular social media platform in China, with 57% of China's total Weibo users and 87% of China's Weibo activities. In 2018, Sina Weibo's monthly active users reached 462 million (2018 Sina Weibo User Development Report). Weibo data have been used in many areas, including public health [25], environmental issues [26], natural disasters [27], and urban land use [28]. Based on the timestamp and location record

of each Weibo user's posts, the user's movement trajectory can be constructed. The resulting intercity migration record can support this study.

The current research based on SFT is based mainly on LBS data provided by Baidu and Tencent in 2015 and 2016 [23,29,30]. Therefore, we use the Sina Weibo data obtained through the API provided by the Sina Weibo platform in 2015 to conduct research, which is convenient for comparison with other related research, published urban economic statistics, and China's 13th Five-Year Development Plan. The Chinese Lunar New Year in 2015 is on February 19. Traditionally, the preparations and celebrations of Spring Festival start on Xiao Nian Day (approximately one week before Spring Festival) and end at the Spring Lantern Festival (two weeks after the Chinese Lunar New Year day). As observed from the Tencent and Baidu LBS data, population flows usually reach a trough on the third day after the Spring Festival, and the population intercity flow before and after this day is reversed in direction and highly symmetrical in magnitude [23,30]. Therefore, the typical SFT period in 2015 includes two periods: February 7–February 21 is the leaving period (two weeks), when the majority of migrants leave the city where they work and go back to their hometown for family reunion, and February 22–March 7 is the return period (two weeks), when people return to the cities where they live after celebrating Spring Festival. In addition, March 8–March 21 (two weeks) is used as the ordinary period (two weeks) for comparison. For these three periods, we obtain a total of 18,152,016 Sina Weibo posts for 360 cities, including four municipalities, 293 prefecture-level cities, some county-level cities, Hong Kong, Macao, and Taiwan, which comprise the majority of China, with a fixed download frequency and uniform geographical distribution [31]. Figure 1a shows the geographical distribution of Sina Weibo posts. Although we can obtain only a small part of Sina Weibo's records for academic research, due to data acquisition restrictions, the number of Sina Weibo posts in each city has a high positive correlation with China's urban population (China Statistic Yearbooks 2015), as shown in Figure 1b, proving that the obtained data represent a reasonable spatial sampling.



**Figure 1.** Chinese Sina Weibo data. (a) Geographical distribution of Sina Weibo posts; (b) The correlation between the number of Sina Weibo post and urban population per region.

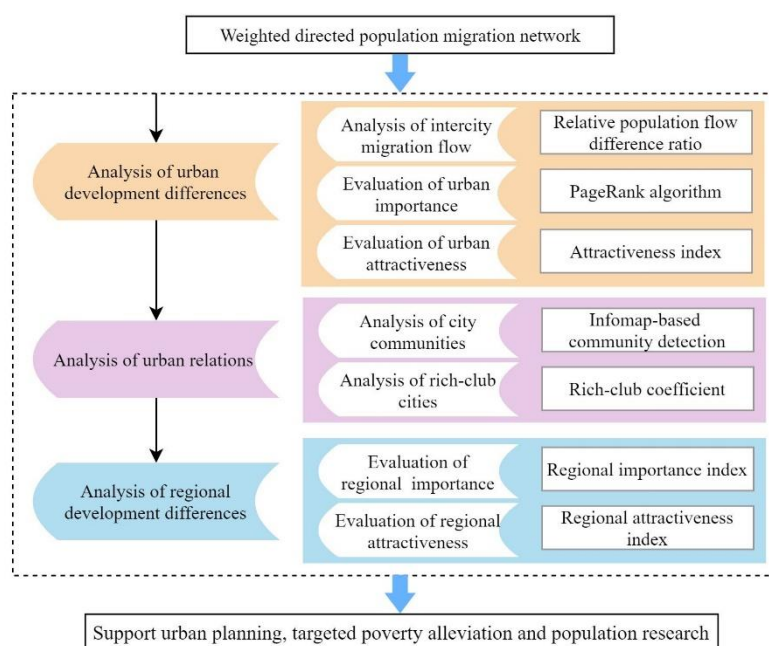
The posts of each Sina Weibo user are extracted from the raw data based on the unique user identity labels. Then, the location information of each Weibo post is spatially joined with the administrative division data of China (from the resource and environment data cloud platform of the Institute of Geographical Sciences and Natural Resource Research of the Chinese Academy of Sciences) through overlay analysis provided by Arcmap10.5.1, to determine the city where the user posted information. According to the timestamp, the location record of each Sina Weibo user can be obtained. Then, it is determined whether the position records of the two sequential posts are made in the same city, and if there is a city change, an intercity movement is recorded. Finally, all of the intercity movement records of the same period are summarized to construct a weighted directed population flow network, with the cities as network nodes and the direction and volume of population flow as the directions and weights of the network edges. Compared with LBS migration data, Baidu migration data count population migration by hourly granularity, which can easily cut off unfinished journeys and trigger an increase in

the number of short trips, while Tencent migration data count population migration data on a daily basis, ignoring night traffic. As personal life updates are one of the major motivations for users to post information, the spatiotemporal behaviors triggered by Spring Festival become an incentive for Weibo users to record their lives. Therefore, the method of measuring intercity migration based on the location changes of Sina Weibo users' posts is flexible and feasible. Based on the movement trajectory of the Sina Weibo users, we constructed the weighted directed networks of intercity population migration flows considering the posts made during the entire SFT period, the leaving period, the returning period, and the ordinary period and obtained 996,901, 306,795, 429,907, and 175,911 intercity population flow records, respectively.

### 3. Methodology

#### 3.1. Overall Methodological Framework

This article proposes a geospatial network analytical framework to analyze the urban/regional development imbalance based on the PMN from the perspectives of geospatial science and complex network science. Three major steps are taken to complete the analysis, as shown in Figure 2. Firstly, the urban development difference is explored based on the intercity population flow difference ratio, PageRank algorithm, and attractiveness index. Secondly, the community detection method and rich-club coefficient are applied to further observe the spatial interactions between cities. Finally, the regional importance index and attractiveness index are used to reveal the regional development imbalance. The methods are described in detail in the following sections.



**Figure 2.** The methodological framework of the geospatial network analysis.

#### 3.2. Attractiveness Index Based on Intercity Migration

Regarding the SFT period, during the leaving period, the main flow of people leave their current residence and go to their hometown for the family reunion, and in the returning period, these migrants return to where they work. This migration phenomenon, triggered by Chinese traditional culture, leads to an imbalance in the population flow between cities in the two periods. Cities with relatively high outflows during the leaving period and relatively high inflows during the returning period are the more developed cities that offer higher salaries and better living quality and absorb more laborers than the less-developed cities. In contrast, cities with relatively high inflows during the leaving period

and relatively high outflows during the returning period are the relatively underdeveloped cities that export laborers. The index of the relative flow difference ratio (*RFDR*) between the population outflow volume and population inflow volume, which is the ratio of the difference between the inflow volume and outflow volume to the total flow volume, is used to describe the unbalanced population flow of a city:

$$RFDR = \frac{num_{inflow} - num_{outflow}}{num_{inflow} + num_{outflow}}, \quad (1)$$

A positive *RFDR* indicates that the inflow volume exceeds the outflow volume and vice versa. When the population migration in a city is balanced during a period, the *RFDR* is approximately zero.

The index of attractiveness is designed based on the expectation that people usually migrate to a more attractive city from a less attractive city. J. Xu et al. (2017) defined attractiveness as the difference in the *RFDR* between the returning period and the leaving period, ignoring the situation in which the *RFDR* is positive or negative in both periods [29]. Whether the *RFDR* is positive or negative in both periods, indicating that the population continues to flow in or out, shows whether a city is more attractive or less attractive. Therefore, here we consider two cases. If the *RFDR* is positive or negative in both periods, we sum to express the attractiveness of the city; otherwise, we use the difference of *RFDR* between the two periods:

$$Attractiveness = \begin{cases} RFDR_{return} - RFDR_{leave}, & \text{if } RFDR_{return} \times RFDR_{leave} \leq 0 \\ RFDR_{return} + RFDR_{leave}, & \text{if } RFDR_{return} \times RFDR_{leave} > 0 \end{cases}, \quad (2)$$

If there are more people entering than leaving a city in both periods, then the attractiveness of the city is positive, and vice versa. Moreover, if there are more people leaving in the leaving period and more people entering in the returning period, then the attractiveness is positive, and vice versa.

### 3.3. Importance Evaluation Based on the PageRank Algorithm

The PageRank algorithm is a commonly used webpage ranking algorithm. It has been widely used in social network analysis [32], transportation planning [33], and network security [34]. The PageRank algorithm takes the number and quality of hyperlinks between web pages as the main factors to analyze the importance of a web page [35]. The basic assumption is that more important pages are often referenced by other web pages. The link from page A to page B is regarded as "page A votes for page B", and the importance of the page is determined based on the source and the importance of the voters. The SFT network is analogous to the internet in that more important cities in the migration network receive or transfer more routes and populations from other cities. Hence, the PageRank algorithm is suitable for understanding the importance of cities in the SFT network. The equation of the PageRank algorithm is:

$$PR(x) = \frac{(1 - \sigma)}{n} + \sigma \sum_{i \in P_x} \frac{PR(Y_i)}{C_{out}(Y_i)}, \quad (3)$$

where  $PR(x)$  is the PageRank value of page  $x$ ,  $P_x$  depicts the set of pages that link to page  $x$ ,  $PR(Y_i)$  is the PageRank value of page  $Y_i$  linked to page  $x$ ,  $C_{out}(Y_i)$  describes the number of links out from page  $Y_i$ ,  $n$  depicts the number of pages, and  $\sigma$  is the damping factor used to deal with pages that have no external links. These pages are considered to be linked to all pages in the network, and the PageRank values of such pages are divided equally among all pages. The values of PageRank can be approximated with high accuracy through several iterations. The more links and the higher the PageRank value of the source page to page  $x$ , the larger the PageRank value of page  $x$  will be.

In the SFT network,  $PR(x)$  is the PageRank value of city  $x$ ;  $P_x$  describes the cities that export travelers to city  $x$ ;  $C_{out}(Y_i)$  expresses the number of links from the city  $Y_i$ , which is weighted by the volume of travelers; and  $n$  depicts the number of cities. Considering that China's transportation network is very developed and the traffic between cities is convenient, we weaken the weight of the damping factor and set it to 0.95. The PageRank algorithm is implemented by the python-NetworkX package.

### 3.4. Community Detection

As a typical feature of a complex network, community structure describes the characteristic that the connections within communities are relatively dense, while connections among the communities are relatively sparse. Recognizing the network communities is conducive to disclosing the latent relations between the network nodes. Community detection methods, such as the label propagation algorithm [36,37] and modularity-based algorithm [38,39], are designed based on the characteristics of different network types (e.g., topology network, binary network, and directed network) and widely used in various fields (e.g., social network analysis, biological network analysis, and internet network analysis). Intercity PMN is a typical weighted directed network, and we can apply appropriate community detection methods to explore the interactions between cities. The Infomap algorithm [40,41] has been shown to perform well in weighted directed networks, especially small networks (less than 1000 nodes), through a variety of comparative tests [42,43] and is suitable for observing the community structure of the intercity SFT network (360 nodes).

Based on information theory [44], the Infomap algorithm is dedicated to identifying network communities by selecting the fewest bits to express the route generated by random walks in the network. Each node is encoded as the Huffman code [45], and a node with a higher access frequency is assigned a shorter code. To divide nodes into different clusters, a two-level description strategy is adopted. Each cluster is given a unique name, and the nodes within each cluster are named with different Huffman codes, which can be reused in different clusters. Differently from a single-layer structure that does not consider community structure, a two-level description strategy allows the nodes within one cluster to share the same code, so that the nodes themselves can be depicted with fewer bit codes, and shorter average bits can be achieved. Regarding a community partition  $M$  of  $n$  nodes into  $m$  clusters, the average description length of a single step  $L(M)$  is:

$$L(M) = q_{\curvearrowright} H(Q) + \sum_{i=1}^m p_{\cup}^i H(P^i), \quad (4)$$

where  $H(Q)$  describes the information entropy of the clusters names;  $q_{\curvearrowright}$  describes the probability that the random walk switches clusters [44];  $H(P^i)$  depicts the information entropy of the movements within cluster  $i$ , including the exit code (a virtual node within each cluster expressing that the random walk is leaving the current cluster) for cluster  $i$ ; and  $p_{\cup}^i$  depicts the probability of movements within cluster  $i$  (including the probability of exiting cluster  $i$ ). The access probability of each node is calculated by the "random surfer" method, which is similar to the PageRank algorithm, and the possible partitions are explored based on a simulated annealing approach [46] and deterministic greedy search algorithm [47]. In this paper, the Infomap algorithm is used to explore the community structure of the SFT network and the ordinary travel network separately, using the R-igraph package [48].

### 3.5. Rich-Club Coefficient

The rich-club coefficient is applied to observe the spatial interactions between the prominent cities. As early as the end of the 19th century, inspired by the social and economic disparity among people in different societies and countries, the 80/20 rule was proposed to describe the phenomenon that a select minority of elements are responsible for the vast majority of the observed outcomes in many real-world settings. Then, the rich-club phenomenon is defined to describe the tendency of prominent elements to establish stronger interactions among themselves than would be expected by random chance. This phenomenon has proven to be applicable in many areas, including transportation networks [49], scientific collaboration networks, and interbank networks [50]. The rich-club coefficient was first defined based on the topology network [51]. Subsequently, Opsahl, Colizza, Panzarasa, and Ramasco (2008) improved the rich-club coefficient  $\varphi_w(r)$  to make it suitable for a weighted directed network [52]. All nodes in the network are ranked in terms of the rich parameter  $r$ . For each value of  $r$ , the nodes whose richness is larger than  $r$  construct the nodes club.  $E_{>r}$  depicts the number of edges connecting the members whose richness is larger than  $r$ ,  $W_{>r}$  sums up the weights of these edges, and  $W_{l,rank}$  is the  $l$ th ranked weight on the edges of the network. Then, we have  $\varphi_w(r)$ :

$$\varphi_w(r) = W_{>r} / \sum_{l=1}^{E_{>r}} W_{l,rank} \quad (5)$$

In addition, the rich-club coefficient  $\varphi_{w,null}(r)$ , which is obtained from the corresponding null model that is random but still comparable to the real network, is introduced as a benchmark for comparison [49]. Therefore, the rich-club effect is determined as follows:

$$\rho_w(r) = \frac{\varphi_w(r)}{\varphi_{w,null}(r)}, \quad (6)$$

Only when the rich-club coefficient of the actual network  $\varphi_w(r)$  is greater than that of its corresponding randomized network  $\varphi_{w,null}(r)$ —that is, only when the ratio  $\rho_w(r)$  is greater than 1—can it be proved that there is a rich-club phenomenon in the network. For the weighted directed network, the richness parameter can be the out-degree (the number of links going out from a node) or out-strength (the sum of the weights attached to these links). High out-degree and out-strength prove the importance of a node because of its high external links and external participation in the network. In the SFT network, the rich-club ratio with these two richness parameters is employed to observe the spatial interaction pattern between the prominent cities using the R-tnet package [53].

### 3.6. Community Evaluation

To further explore the regional differences in China's development, we describe the regional differences based on the attractiveness and importance of the cities that make up the communities. The mean of the PageRank values of the cities included in the community is defined to describe the importance of the community:

$$PR(C_n) = \sum_{i \in C_n} PR(city_i) / Num_{C_n}, \quad (7)$$

where  $PR(city_i)$  depicts the PageRank value of city  $i$ ,  $C_n$  expresses the city subgroup included in community  $n$ ,  $Num_{C_n}$  is the number of cities included in community  $n$ , and  $PR(C_n)$  is the PageRank value of community  $n$ . If the cities in the community have relatively high PageRank values, then the community's PageRank value will be relatively high, and vice versa.

Concerning the attractiveness of the community, urban attractiveness is calculated based on the relative flow difference ratio of the inflow and outflow of the urban population, so the average city attractiveness is not suitable to quantify the attractiveness of the entire community. Concerning a more prominent city, its attractiveness should have a more significant impact on the overall attraction of the community. Therefore, the ratio of the city PageRank value to the maximum PageRank value is used as the weight of urban attractiveness on the attractiveness of the community. The average weighted urban attractiveness within the community is used to define the attractiveness of the community:

$$Attractiveness_{C_n} = \sum_{i \in C_n} \frac{PR(city_i)}{Max_{PR}(city)} \times Attractiveness_{city_i} / Num_{C_n}, \quad (8)$$

where  $Max_{PR}(city)$  is the maximum PageRank value of the cities and  $Attractiveness_{city_i}$  is the *Attractiveness* of city  $i$ . The more influential and attractive the cities in the city community are, the more attractive the city community is.

## 4. Results

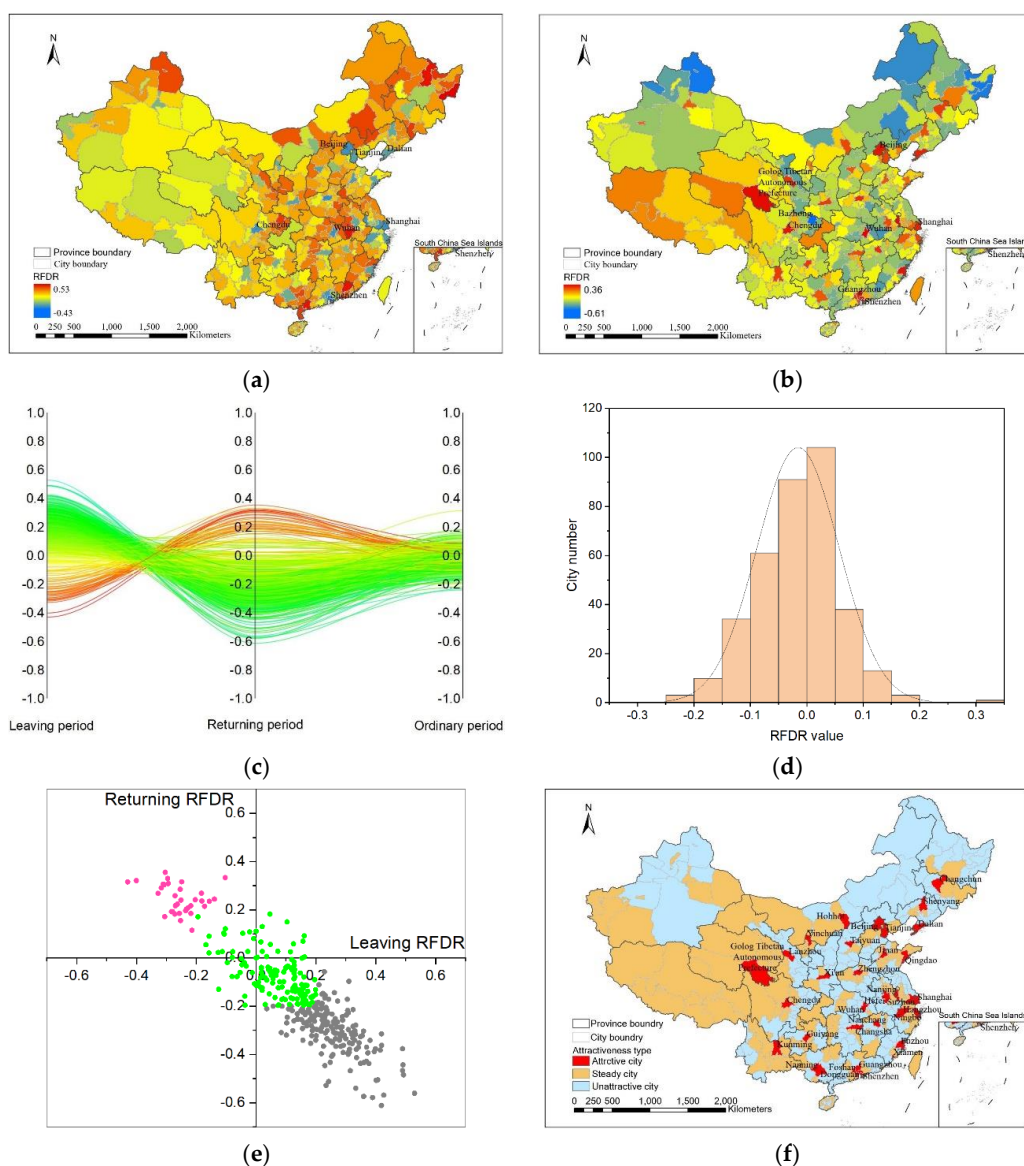
### 4.1. Imbalance of Urban Development

#### 4.1.1. The Difference in RFDR

Figure 3 displays the RFDR of Chinese cities in the three periods, and an opposite trend is found during the leaving period and returning period. In the leaving period, as shown in Figure 3a, high



positive values appear mainly in central, northeast, and northwest China, with Jixi in Heilongjiang Province having the highest value of 0.53. Beijing, Dalian, Tianjin, the Yangtze River Delta, and the Pearl River Delta show significant negative values, with Shenzhen of Guangdong Province showing the lowest value of  $-0.43$ . In the returning period (Figure 3b), significant positive values appear in Beijing, the Pearl River Delta, Dalian, the Yangtze River Delta, Guoluo, Chongqing, and some cities in western China. Wuhan in Hubei Province has the highest value of 0.36. Negative values are mainly observed in northeast, northwest, and central China, with Bazhong in Sichuan Province having the lowest value of  $-0.61$ . A strong negative correlation of *RFDR* between the leaving period and the returning periods is presented in Figure 3c. A city that has more relative population inflow in the leaving period has more relative population outflow in the returning period, and vice versa. During the ordinary period, the *RFDR* distribution roughly shows a Gaussian distribution  $N(-0.015, 0.005)$ , with 98% of the values distributed between  $-0.2$  and  $0.2$  (Figure 3d).

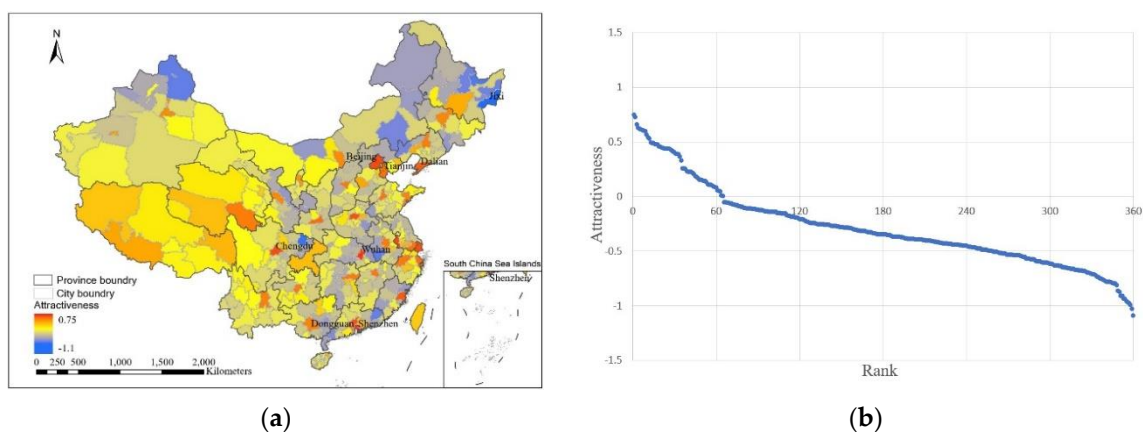


**Figure 3.** Relative flow difference ratio (*RFDR*) of cities in different periods. (a) Leaving period; (b) Returning period; (c) *RFDR* change of cities in the three periods; (d) *RFDR* distribution in the ordinary period; (e) City classification based on *RFDR* change; (f) Geographical distribution of different types of cities.

It can be seen from the variation in *RFDR* in three periods that the tradition of family reunion at Spring Festival caused a strong opposite population flow trend before and after the Spring Festival, and the trend of population flow becomes stable during the ordinary period. Therefore, according to the *RFDR* distribution during the ordinary period, we use the interval  $(-0.2, 0.2)$  as the stability interval and judge a city's attractiveness based on the change in the *RFDR* of the city in the leaving and returning periods (Figure 3e). Thirty-four cities are classified as attractive cities because of the apparent negative value and positive value of the *RFDR* in the leaving and returning period, respectively, which indicates that 10% of Chinese cities have significant appeal to migrant workers. In total, 145 cities are regarded as stable cities because the *RFDR* values in both periods remain within the stable interval. Finally, 181 cities are rated as unattractive cities because of the prominent positive value and negative value of the *RFDR* in the leaving and returning periods, respectively. Similarly, Long and Wu (2016) identified 180 shrinking cities in China, with a decline in population density, based on China's censuses in 2000 and 2010 [54]. As shown in Figure 3f, three distinct attractive urban clusters are located in the Yangtze River Delta, Pearl River Delta, and Beijing-Tianjin-Hebei region. In addition, most provinces, except Xinjiang, Tibet, Taiwan, Hainan, and Heilongjiang, contain one or two attractive cities, surrounded by several stable cities and unattractive cities. In comparison with these five regions, Xinjiang, Hainan, and Heilongjiang contain several unattractive cities, reflecting the phenomenon of population loss in some subregions, while the population changes in Tibet and Taiwan are stable.

#### 4.1.2. The Difference in Urban Attractiveness

The spatial distribution of urban attractiveness appears to be similar to that of the *RFDR* in the returning period (Figure 4a). Negative values appear mostly in northeast, northwest, and central China, with Jixi of Heilongjiang Province having the lowest value of  $-1.1$ . Higher values appear in Beijing, Dalian, Chengdu, Tianjin, the Pearl River Delta, and the Yangtze River Delta. Shenzhen yields the highest value of  $0.75$ , followed by Dongguan and Wuhan. Figure 4b describes the distribution of the attractiveness value. Only 64 cities (17.8%) show positive attractiveness, while most cities have negative attractiveness, reflecting the severe imbalance in urban development in China. A small number of influential cities are relatively developed and absorb labor from underdeveloped regions, while most cities are relatively backward in development and export labor to these developed cities.

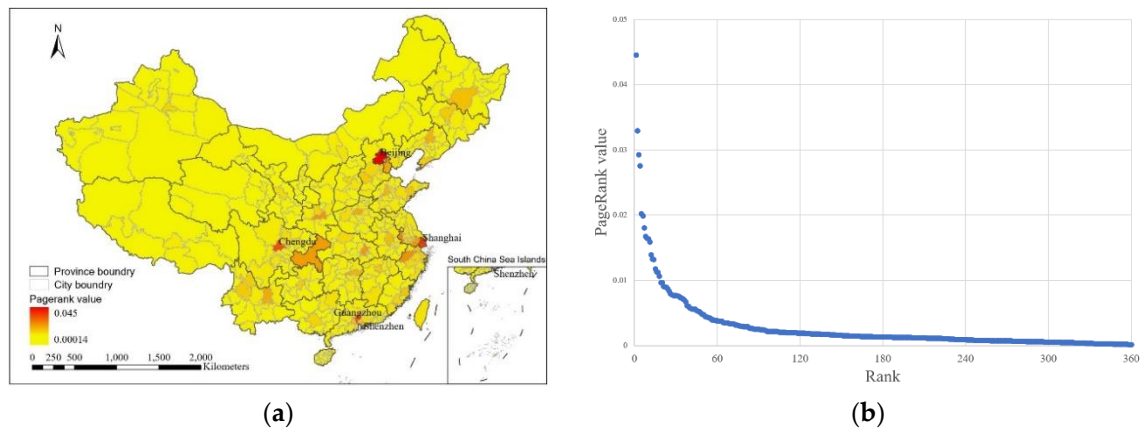


**Figure 4.** The attractiveness of the Chinese cities. (a) Spatial distribution of urban attractiveness; (b) Attractiveness distribution.

#### 4.1.3. The Difference in Urban Importance Based on PageRank

The importance of cities is similar to their classification and attractiveness (Figure 5a). The attractive cities classified based on the change in *RFDR* have relatively higher PageRank values, with Beijing yielding the highest value of  $0.045$ , followed by Shanghai, Chengdu, Guangzhou, and Shenzhen. Moreover, lower values appear mostly in the western and northern regions of China. Figure 5b shows

a heavy-tailed distribution of the PageRank values of cities, reflecting significant differences between Chinese cities. For a few core cities, the higher the ranking is, the higher the PageRank value, and there are significant differences between cities. However, for most of the remaining cities, their PageRank values are low, and there is little difference between cities. This is somewhat similar to the 80/20 rule: a few core Chinese cities have greater influence, while most other cities lack core competitiveness.

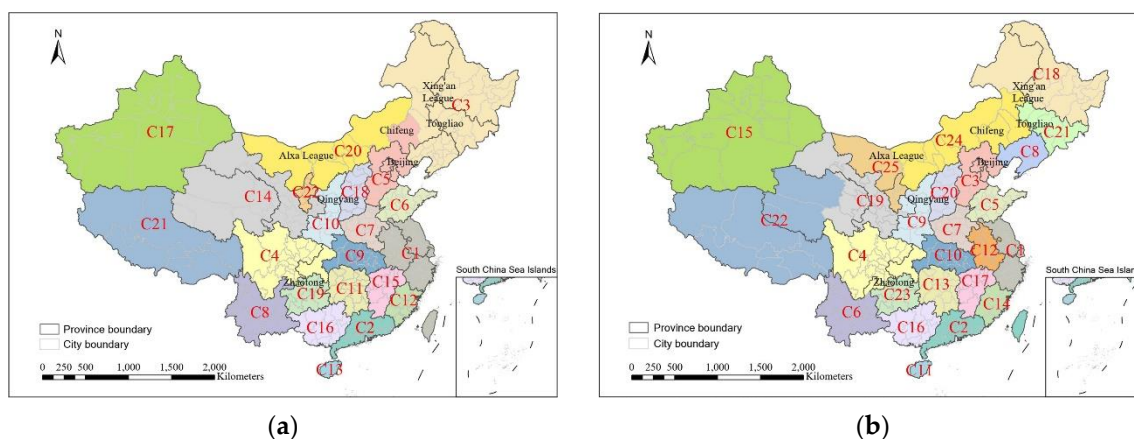


**Figure 5.** Importance of Chinese cities. (a) Spatial distribution of urban importance based on PageRank; (b) PageRank value distribution.

#### 4.2. Intercity Interaction Analysis Based on Population Flow

##### 4.2.1. City Communities Based on the Intercity Migration Network

Figure 6 describes the community structures of cities during the SFT period and the ordinary period and their mapping relationship. For the ordinary period, 25 communities of cities are detected, as shown in Figure 6b. Most of the community divisions are consistent with the Chinese administrative division of the province, while some communities demonstrate two phenomena of interprovincial aggregation and fragmentation. Regarding interprovincial aggregation, community C1 (Yangtze River Delta) consists of Shanghai, Jiangsu Province, and Zhejiang Province; community C2 (Pearl River Delta) consists of Hong Kong, Macao, and Guangdong Province; community C3 (Beijing-Tianjin-Hebei region) consists of Beijing, Tianjin, and Hebei Province; and community C4 consists of Chongqing, Sichuan Province, and a city from an adjacent south province. These four regions have many cities with a high PageRank and attractiveness, forming tight spatial organizations based on developed infrastructure, such as transportation and communication. By contrast, the northeastern and western parts of Inner Mongolia join Heilongjiang Province and Ningxia Province to form communities C18 and C25, respectively, and the eastern and western parts of Qinghai Province are clustered with Gansu Province and Tibet to form communities C19 and C22. This segmentation phenomenon is affected by many factors, such as the economy, culture, and geography. For example, both western Qinghai and Tibet are mainly Tibetan living areas [55], while eastern Inner Mongolia and Heilongjiang Province have similar climates and living customs [56]. The community structure of cities during ordinary time reflects that the majority of the trips are intraprovincial trips and interprovincial trips between neighboring provinces.

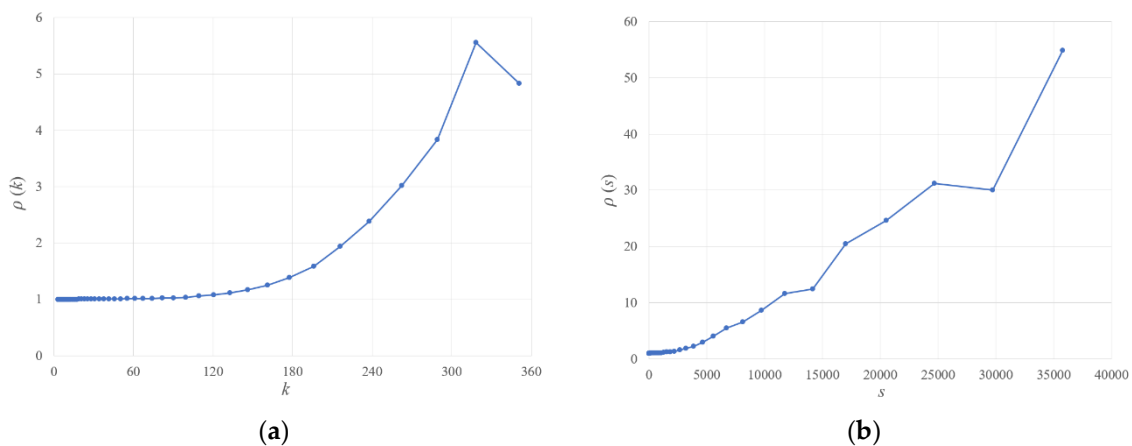


**Figure 6.** City communities. (a) Spring Festival travel (SFT) period; (b) Ordinary period.

The community structure of cities during the SFT period is similar to that during the ordinary period, except that some communities merge and several cities move between neighboring communities, with only 22 communities remaining (Figure 6a). For the merger of the communities, the northeastern part of Heilongjiang Province, Liaoning Province, Jilin Province, and Inner Mongolia's eastern cities constitute Community 3 in the northeastern part of China; Anhui Province, Zhejiang Province, Jiangsu Province, and Shanghai in eastern China constitute Community C1. The cities of Alxa League, Zhaotong, Qingyang, and western Qinghai, which are separated by neighboring provinces during the ordinary period, re-establish communities C20, C8, and C14 with the cities belonging to the same province during the SFT period. However, differently from these cities, during the SFT period, Chifeng city detaches from its province and joins the neighboring community of Beijing, Tianjin, and Hebei Province. This phenomenon is mainly affected by the family reunion behavior triggered by the traditional culture of the Spring Festival. The increase in the size of some communities also reflects the increase in long-distance travel during the SFT period.

#### 4.2.2. Rich-Club Effect

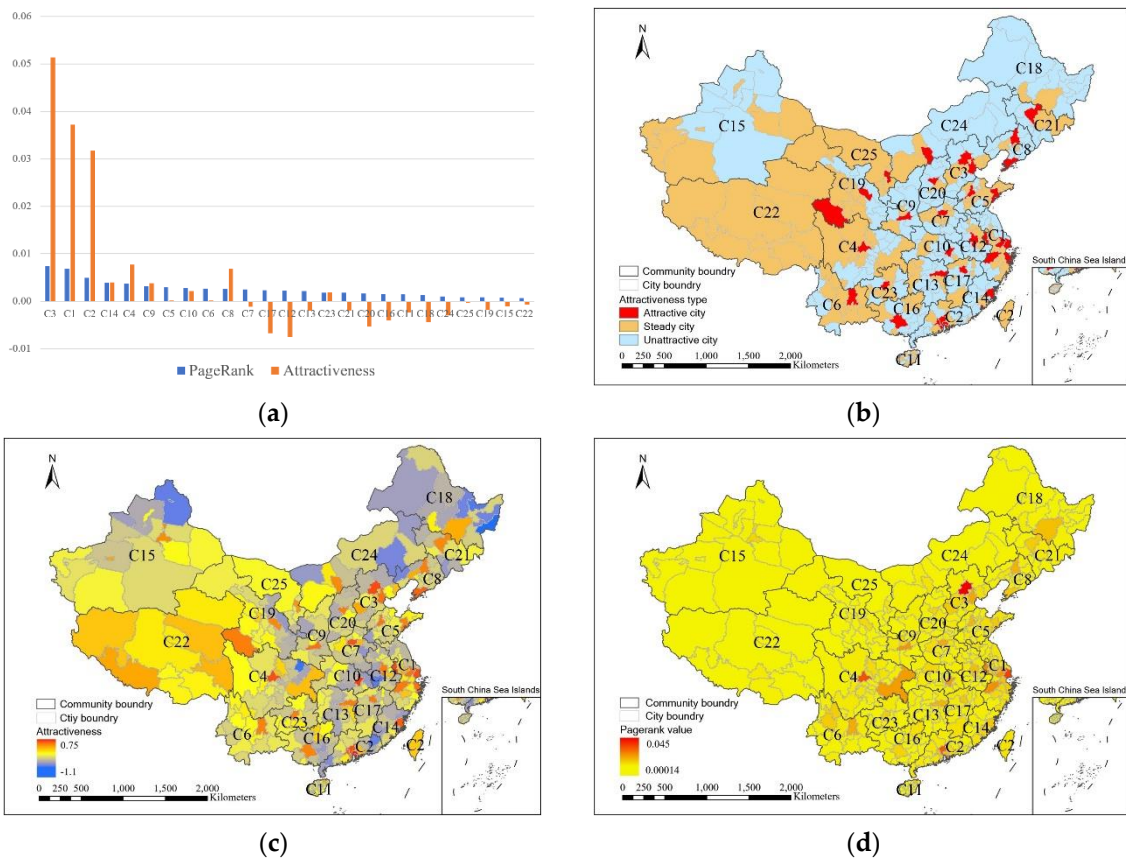
Regarding the rich-club coefficient, Figure 7 shows that the vast majority of  $\rho_w(k)$  and  $\rho_w(s)$  are greater than 1 and present an upward trend, reflecting the remarkable rich-club phenomenon in the SFT network. This means that the prominent cities in China tend to engage in stronger interactions among themselves. In addition, a significant “demarcation point” feature can be found from the change in the curves of both  $\rho_w(k)$  and  $\rho_w(s)$ . Regarding  $\rho_w(k)$ ,  $k = 330$  is a demarcation point: for points less than 330, the curve grows steadily, and for points larger than 330, the curve decreases. For  $\rho_w(s)$ ,  $s = 24,687$  is a demarcation point because the curve basically continues to rise, but there is a downward wave after this point. The prominent rich-club cities are extracted based on these demarcation points. Using  $k > 330$  as the selection criterion, nine cities are obtained—namely, Beijing, Shenzhen, Chengdu, Shanghai, Chongqing, Sanya, Guangzhou, Hangzhou, and Wuhan. Each of these cities has a migration connection with more than 90% of China's cities. With  $s > 24,687$  as the selection threshold, five cities are selected, including Beijing, Shanghai, Chengdu, Guangzhou, and Shenzhen. The analysis based on  $\rho_w(k)$  emphasizes the topological connection between cities, while the analysis based on  $\rho_w(s)$  focuses more on the strength of association between cities. Based on different selection criteria for the rich parameter, we can obtain rich-club cities at different rich levels. However, the core rich-club cities should have the most connections and the greatest connection strength. Therefore, we select the intersection of the two city lists as the core rich-club members, namely, Beijing, Shanghai, Chengdu, Guangzhou, and Chongqing. The population migration flow involving these five rich-club members accounts for 23.2% of the total migration flow during the SFT period. Located in different city communities, these core rich-club cities form the backbone network of the SFT network and can promote cross-regional population flow through close interconnections.



**Figure 7.** Rich-club coefficients in the SFT network. (a)  $r=k$  (out-degree); (b)  $r=s$  (out-strength).

#### 4.3. Imbalance of Regional Development

To further understand the regional development differences in China, we use the urban community structure of the intercity migration network in the ordinary period as an embodiment of the urban agglomeration to analyze the importance and attractiveness of the regions (Figure 8). Figure 8a shows the PageRank and attractiveness of communities. C1 (Yangtze River Delta), C2 (Pearl River Delta), and C3 (Beijing-Tianjin-Hebei region) are the three most important and most attractive city communities. Moreover, the southeastern coastal areas of China, including C8, C3, C5, C1, C14, and C2, have high PageRank values and positive attractiveness values. In contrast, inland communities adjacent to these communities, including C18, C21, C24, C20, C7, C12, C17, C13, and C16, show relatively significant negative attractiveness values, regardless of their PageRank values. The four regions distributed in western and northwestern China (C25, C19, C15, and C22) exert the least influence, but the population flow in these regions is relatively stable and does not show significant negative attractiveness. The remaining five communities (C4, C6, C9, C10, and C23), clustered in central and southwestern China, have relatively high PageRank values and positive attractiveness values.

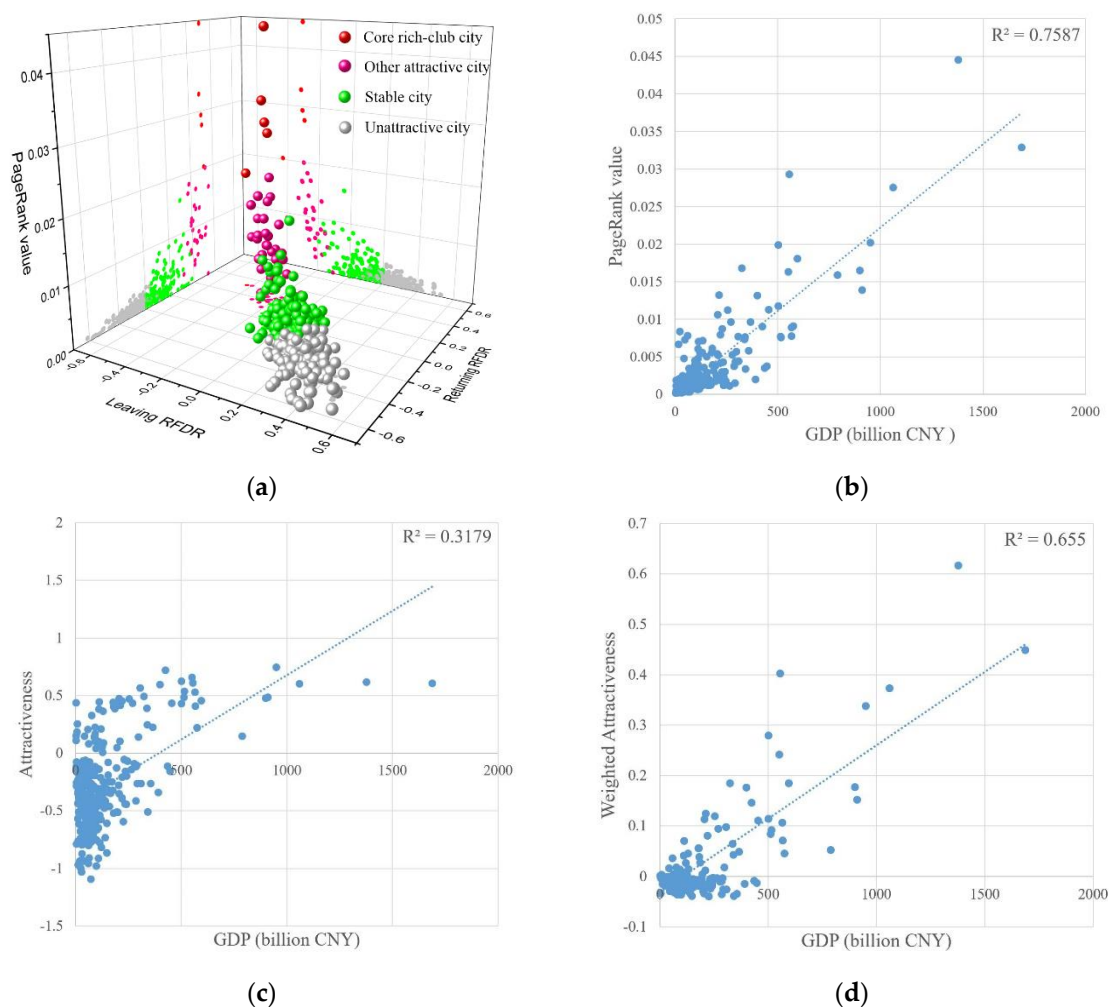


**Figure 8.** Community evaluation. (a) The PageRank value and attractiveness of communities, sorted by PageRank value; (b) Overlay of city types and city communities; (c) Overlay of the attractiveness values of cities and city communities; (d) Overlay of PageRank values of cities and city communities.

## 5. Discussion

### 5.1. Indices of Urban Development Based on PMN

There is a clear correlation between the PMN-based urban development indices, as shown in Figure 9a. The relationship between PageRank and *RFDR* presents a trend in which the PageRank value decreases while the city's *RFDR* value goes from low to high during the leaving period and from high to low during the returning period. Additionally, the attractive cities have relatively high PageRank values; correspondingly, the unattractive cities have relatively low PageRank values. It can be inferred that cities with higher PageRank values are more competitive, thus showing significant and opposite *RFDR* values during the two periods of Spring Festival. In addition, the core rich-club cities show the top five PageRank values and have significantly negative *RFDR* values in the leaving period and significantly positive *RFDR* values in the returning period, reflecting the prominent influence of these cities in the PMN.



**Figure 9.** The relationships between indices based on SFT data and urban economic indices. (a) PageRank–RFDR; (b) PageRank–GDP; (c) Attractiveness–GDP; (d) Weighted attractiveness–GDP.

PMN-based urban development indices are highly correlated with urban economic development levels. To find the relation between the indices used in this paper and urban development, the PageRank values, attractiveness values, and weighted attractiveness values (the attractiveness value multiplied by the ratio of the PageRank value with the cities’ maximum PageRank value) calculated from the migration network are compared with the urban GDP from the China Statistical Yearbook 2015. The correlation coefficient between the PageRank value and GDP yields a very high value (Figure 9b), indicating that the PageRank value based on the migration network can reflect the development level of a city. Although attractiveness and weighted attractiveness are both positively correlated with GDP, the correlation coefficient between weighted attractiveness and urban GDP is higher than that between attractiveness and GDP. This is reasonable because the attractiveness value reflects the relative attractiveness of the city itself based on the relative population flow ratio and is not very suitable for the horizontal comparison of development differences between cities. Weighted attractiveness considers the difference in the level of development between cities and can better reflect the difference in attraction between the cities, thus having a higher correlation with urban GDP. This also shows that using cities’ weighted attractiveness to analyze the attraction of communities is more suitable than using attractiveness.

However, although the PageRank value has a higher correlation with urban GDP than attractiveness and weighted attractiveness, having a high PageRank value does not mean that a city also has lower and higher RFDR values in the leaving period and returning period, respectively (Figure 9c,d). This is

because the attractiveness of a city depends on not only its level of development but also various natural and social factors, especially the development levels of surrounding cities. For example, consider two cities with similar development situations, one adjacent to more developed cities and the other to underdeveloped cities. Then, influenced by the surrounding cities, the former city has a relatively low attractiveness, and the latter city has a relatively high attractiveness. This rule also applies to the city community. The PageRank value of C14 is slightly higher than that of C4, but C4 is almost twice as attractive as C14, largely because C14's attractiveness is affected by the adjacent C2 and C1 with higher PageRank values, and the PageRank values of the communities around C4 are relatively lower.

### 5.2. Rich-Club Cities with Interregional Connections

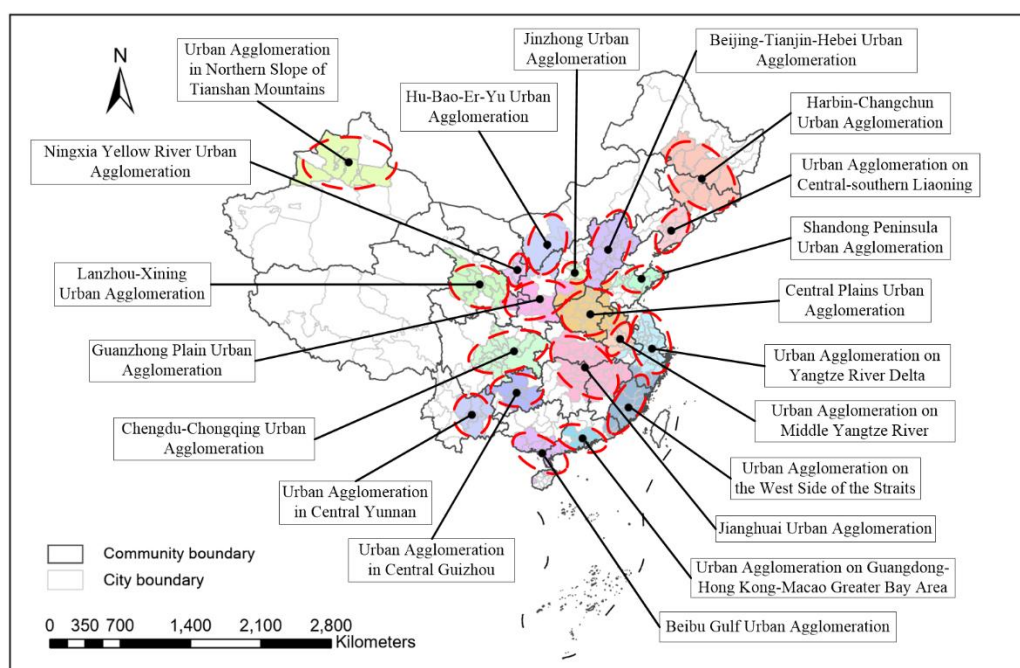
The rich-club coefficient based on two rich parameters (connections and weights) proves that there is a rich-club phenomenon in the SFT network, which means that these influential cities are not isolated from each other but have a close interaction. Although this finding is similar to the finding of a study based on the 2016 Tencent LBS data [30], there are two subtle differences that require further discussion. In the PMN based on the Tencent LBS data, which includes 369 cities, the maximum out-degree of the city nodes is just 272, and the breakpoint of  $k = 200$  is selected to determine 11 rich-club cities based on their connections. However, in the PMN based on the Sina Weibo data, we find better connectivity between cities, with the maximum out-degree of the urban nodes being 353, and nine cities connecting with more than 330 cities. Additionally, the six core rich-club cities obtained based on the Tencent LBS data are all located in the coastal areas (four of which are consistent with our results) but ignore the central hub of Chengdu. Moreover, the connections involving these six core rich-club cities account for 49.57% of the population flow, more than twice the 23.3% of movements involved in the five core rich-club cities in this study. This inconsistency is mostly due to data limitations. As shown in [23], the Tencent LBS data provide only the top 10 inflows and top 10 outflows per city per day. Therefore, the resulting PMN cannot adequately describe the intercity Spring Festival migration. On the other hand, the flow data provided by the Tencent LBS data lack passenger information and cannot integrate the segments of a traveler's trip into the entire travel route and thus are unable to describe long-distance travel.

The community structure of the cities demonstrates that the migration between cities is influenced by geographical proximity, and the cities within the community are spatially adjacent. Moreover, five core rich-club cities are scattered in the north, central, eastern, and southern parts of China and serve as transportation hubs, connecting various scattered areas through close interconnections. Additionally, these core rich-club cities are attractive cities and have the highest PageRank values, meaning that these cities have greater attractiveness and higher competitiveness. It is inferred that a small number of rich-club cities dominated by Beijing, Shanghai, Chengdu, Guangzhou, and Shenzhen serve as critical regional nodes for local economic development, integrating spatially dispersed areas and promoting effective interaction across the country with close interconnection. It can be seen from China's development plan for the transportation system (China's 13th Five-Year Plan for the development of a modern comprehensive transport system) that China is currently systematically building multilevel transportation hubs to promote interconnectivity across the country, and the five core rich-club cities identified in this paper—namely, Beijing, Shanghai, Guangzhou, Shenzhen, and Chengdu—are listed as the first comprehensive international transportation hubs (among a total of seven). It can be inferred that the development of China's transportation network system will further enhance the rich-club characteristics in the population flow. These few rich-club cities form a dense and interconnected backbone network, which attracts, absorbs, and disseminates large-scale population flows.



### 5.3. City Communities with Urban Agglomeration Planning

Figure 10 shows the distribution of urban agglomerations outlined in China's 13th Five-Year Plan (2016–2020). As shown in Figures 3–5 and Figure 8, these planned urban agglomerations include almost all attractive cities and high PageRank cities and cover all city communities except C22. However, among these urban agglomerations, only the development targets of the Yangtze River Delta (C1), the Pearl River Delta (C2), and Beijing-Tianjin-Hebei (C3) are world-class urban agglomerations, while other urban agglomerations are still in the early stages of cultivation or development, which is consistent with our results. C1, C2, and C3, which contain these three urban agglomerations, have the highest PageRank values and significant attractiveness. In 2015, the economic aggregate of these three urban agglomerations accounted for more than 40% of the national economy (China Urban Agglomeration Integration Report 2019).



**Figure 10.** Planned urban agglomerations in China (China's 13th Five-Year Plan).

However, influenced by the eastern coastal areas, the attractiveness of the adjacent northeastern and central communities is significantly negative, indicating a large number of unattractive cities. The cultivation of urban agglomerations in the northeast and central regions will help to promote the coordinated and sustainable development of the local economy. The four city communities in the western region yield the lowest PageRank value, and the corresponding urban agglomeration planning could revitalize local development. It is observed that China's regional development level shows a gradient pattern of high to low from east to west, which is consistent with the results of studies based on DMSP/OLS nighttime light data 1992–2013 [57], the Development and Life Index 2000–2012 [58], and the China City Statistical Yearbook 2010 [59]. It is worth noting that C22 yields the lowest PageRank value, with no attractive city or high PageRank city, which indicates that the overall development level of this region is backward. However, the attractiveness value of C22 is near zero, indicating that there is no significant population loss in this region. This is mainly caused by cultural diversity; despite the low economic level in the Tibet area, the unique cultural traditions and lifestyle habits in the area restrict local population outflows [60]. In addition, in the community structures of the two periods, Taiwan has established communities with the Pearl River Delta and the Yangtze River Delta, indicating population flow between Taiwan and the Pearl River Delta and Yangtze River Delta. However, during these two periods, Taiwan's *RFDR* values were stable and did not show significant

population inflows or outflows. This is mainly attributed to the current cross-strait policy [61]. On the whole, the national urban agglomeration plan in China is reasonable. However, for areas such as Tibet that are economically backward and lack external connections, it is more necessary for the state to carry out targeted poverty alleviation.

## 6. Conclusions

In this study, the social sensing approach based on Sina Weibo data is used to observe the intercity migration flow in China around Spring Festival and reveal the imbalanced development and spatial interactions among cities. Regarding the development of a city, the *RFDR* around Spring Festival indicates the differentiated development of Chinese cities. A small number of cities show significant attractiveness, while nearly half of the cities are unattractive. Then, the analysis based on the PageRank algorithm and the attractiveness index further reveals the differential development of cities. It is worth noting that the attractiveness of a city depends on not only its development level but also the development level of the neighboring cities. This also applies to regional differences. Regarding the spatial interactions between cities, the community structure based on the PMN presents apparent geographical proximity. The interprovincial integration and fragmentation that occur in the ordinary period are influenced by various factors, such as the economy, culture, and geography. By contrast, the phenomenon of community integration and conversion in several cities during the SFT period is mostly driven by the family reunion triggered by Chinese traditional culture. In addition, the analysis based on the rich-club coefficient confirms the existence of the rich-club phenomenon in the intercity PMN. A small number of rich-club cities dominated by Beijing, Shanghai, Chengdu, Guangzhou, and Shenzhen play essential roles in regional development and have made outstanding contributions to promoting cross-regional long-distance migration. The evaluation of the city community reveals China's regional development differences. The eastern coastal areas represented by the Pearl River Delta, Yangtze River Delta, and Beijing-Tianjin-Hebei region are economically developed and show apparent appeal, while the eastern and northeastern regions adjacent to the developed coastal cities reflect significant labor loss. The population flows of the southwestern region are relatively balanced, but the overall economic development level of the western region is low. The results of regional evaluation prove the rationality of the 13th Five-Year Plan regarding the construction of urban agglomerations. However, although Tibet maintains a relatively stable population flow due to its unique culture, its low economic level requires further government support to revitalize local urban development.

This study focuses on the population migration network around the Spring festival. Similarly to the Chinese Spring festival, Thanksgiving Day is a traditional holiday for family reunions that is popular in several countries in North America. The geospatial network analysis method proposed in this paper can also be used to discuss local development differences in North America. Furthermore, this methodological framework can also be applied to other related fields (e.g., social networks and traffic networks) to identify the differences and connections between nodes and subgroups in a network. For example, in a transit-oriented development model, the regional connectivity situation can be explored based on the public transportation network to assist in new transportation planning. This paper shows that population migration can be used for urban/regional development studies. However, urban studies based on crowdsourced geographic information have emerged only recently, and a much longer series and higher spatial-temporal resolution are needed to fully demonstrate the urban development process and the spatial interaction between cities. In the next step, it would be beneficial to go beyond Sina Weibo data to perceive urban/regional development across different types of crowdsourced geospatial data. Besides, for migration research, the attribute information of social media users (e.g., age, gender, and education) and social media content (e.g., text, location, and time) may be used in combination to further explore the reasons (e.g., work or study) for people's migration.

**Author Contributions:** Conceptualization, R.Z. and Y.W.; methodology, R.Z.; software, R.Z. and R.X.; data curation, M.J.; writing—original draft preparation, R.Z.; writing—review and editing, D.L., M.J., J.G., J.Y. and L.M.; visualization, R.Z.; supervision, L.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research is supported by China’s National Key R&D Program “Data Acquisition and Location-based Information Aggregation for Pan-Information Map”, grant number No. 2017YFB0503500 and National Natural Science Foundation of China, grant number No. 41901335. The APC is funded by the German Research Foundation (DFG) and the Technical University of Munich (TUM) in the framework of the Open Access Publishing Program.

**Acknowledgments:** The support provided by the China Scholarship Council (CSC) during the Ph.D. study of ‘Ruoxin Zhu’ in TUM is acknowledged.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

- Combes, P.P.; Lafourcade, M.; Thisse, J.F.; Toutain, J.C. The rise and fall of spatial inequalities in France: A long-run perspective. *Explor. Econ. Hist.* **2011**, *48*, 243–271. [[CrossRef](#)]
- Jeffrey, G. Williamson Regional Inequality and the Process of National Development: A Description of the Patterns. *Econ. Dev. Cult. Change* **1965**, *13*, 1–84. [[CrossRef](#)]
- Perera, L.D.H.; Lee, G.H.Y. Have economic growth and institutional quality contributed to poverty and inequality reduction in Asia? *J. Asian Econ.* **2013**, *27*, 71–86. [[CrossRef](#)]
- Hirschman, A.O. *The Strategy of Economic Development*; Yale University Press: New Haven, CT, USA, 1958.
- Myrdal, G. *Political Element in the Development of Economic Theory*; Transaction Publishers: Piscataway, NJ, USA, 1990; ISBN 0-88738-827-2.
- Perroux, F. Economic Space: Theory and Applications. *Q. J. Econ.* **1950**, *64*, 89–104. [[CrossRef](#)]
- Prebisch, R. Commercial policy in the underdeveloped countries. *Am. Econ. Rev.* **1959**, *49*, 251–273. [[CrossRef](#)]
- Lessmann, C. Spatial inequality and development—Is there an inverted-U relationship? *J. Dev. Econ.* **2014**, *106*, 35–51. [[CrossRef](#)]
- Wang, Y.; Deng, Y.; Ren, F.; Zhu, R.; Wang, P.; Du, T.; Du, Q. Analysing the spatial configuration of urban bus networks based on the geospatial network analysis method. *Cities* **2020**, *96*, 102406. [[CrossRef](#)]
- Wang, Z.; Hu, B. China’s reform and opening-up and international law. *Chinese J. Int. Law* **2010**, *9*, 193–203. [[CrossRef](#)]
- Long, Y.; Zhai, W.; Shen, Y.; Ye, X. Understanding uneven urban expansion with natural cities using open data. *Landsc. Urban Plan.* **2018**, *177*, 281–293. [[CrossRef](#)]
- Feng, Y.; Pu, Y.; Ma, J.; Chen, G.; Wang, J. Research on the change of interprovincial migration flows in China based on network autocorrelation. *J. Northwest Norm. Univ. Natural Sci.* **2017**, *53*, 119–127. [[CrossRef](#)]
- Goodchild, M.F. Citizens as sensors: The world of volunteered geography. *GeoJournal* **2007**, *69*, 211–221. [[CrossRef](#)]
- Shan, J.; Qin, K.; Huang, C.; Hu, X.; Yu, Y.; Hu, Q.; Lin, Z.; Chen, J.; Jia, T. Methods of Crowd Sourcing Geographic Data Processing and Analysis. *Geomatics Inf. Sci. Wuhan Univ.* **2014**, *39*, 390–396. [[CrossRef](#)]
- Finch, K.C.; Snook, K.R.; Duke, C.H.; Fu, K.W.; Tse, Z.T.H.; Adhikari, A.; Fung, I.C.H. Public health implications of social media use during natural disasters, environmental disasters, and other environmental concerns. *Nat. Hazards* **2016**, *83*, 729–760. [[CrossRef](#)]
- Zhang, X.; Li, W.; Zhang, F.; Liu, R.; Du, Z. Identifying urban functional zones using public bicycle rental records and point-of-interest data. *ISPRS Int. J. Geo Inf.* **2018**, *7*, 459. [[CrossRef](#)]
- Griffin, G.P.; Jiao, J. Crowdsourcing Bike Share Station Locations: Evaluating Participation and Placement. *J. Am. Plan. Assoc.* **2019**, *85*, 35–48. [[CrossRef](#)]
- Can, U.; Alatas, B. Big social network data and sustainable economic development. *Sustainability* **2017**, *9*, 2027. [[CrossRef](#)]
- Steed, R.J.; Fuenzalida, A.; Bossu, R.; Bondár, I.; Heinloo, A.; Dupont, A.; Saul, J.; Strollo, A. Crowdsourcing triggers rapid, reliable earthquake locations. *Sci. Adv.* **2019**, *5*, eaau9824. [[CrossRef](#)]
- Zhu, R.; Lin, D.; Jendryke, M.; Zuo, C.; Ding, L.; Meng, L. Geo-tagged social media data-based analytical approach for perceiving impacts of social events. *ISPRS Int. J. Geo Inf.* **2019**, *8*, 15. [[CrossRef](#)]
- Giuffrida, N.; Le Pira, M.; Inturri, G.; Ignaccolo, M. Mapping with stakeholders: An overview of public participatory GIS and VGI in transport decision-making. *ISPRS Int. J. Geo Inf.* **2019**, *8*, 198. [[CrossRef](#)]

22. Liu, Y.; Liu, X.; Gao, S.; Gong, L.; Kang, C.; Zhi, Y.; Chi, G.; Shi, L. Social Sensing: A New Approach to Understanding Our Socioeconomic Environments. *Ann. Assoc. Am. Geogr.* **2015**, *105*, 512–530. [[CrossRef](#)]
23. Li, J.; Ye, Q.; Deng, X.; Liu, Y.; Liu, Y. Spatial-Temporal Analysis on Spring Festival Travel Rush in China Based on Multisource Big Data. *Sustainability* **2016**, *8*, 1184. [[CrossRef](#)]
24. Zhang, L.; Pentina, I. Motivations and usage patterns of weibo. *Cyberpsychology Behav. Soc. Netw.* **2012**, *15*, 312–317. [[CrossRef](#)] [[PubMed](#)]
25. Tian, X.; Batterham, P.; Song, S.; Yao, X.; Yu, G. Characterizing Depression Issues on Sina Weibo. *Int. J. Environ. Res. Public Health* **2018**, *15*, 764. [[CrossRef](#)] [[PubMed](#)]
26. Wang, S.; Paul, M.J.; Dredze, M. Social media as a sensor of air quality and public response in China. *J. Med. Internet Res.* **2015**, *17*, e22. [[CrossRef](#)] [[PubMed](#)]
27. Wang, Y.; Ruan, S.; Wang, T.; Qiao, M. Rapid estimation of an earthquake impact area using a spatial logistic growth model based on social media data. *Int. J. Digit. Earth* **2018**, 1–20. [[CrossRef](#)]
28. Zhang, Y.; Li, Q.; Huang, H.; Wu, W.; Du, X.; Wang, H. The Combined Use of Remote Sensing and Social Sensing Data in Fine-Grained Urban Land Use Mapping: A Case Study in Beijing, China. *Remote Sens.* **2017**, *9*, 865. [[CrossRef](#)]
29. Xu, J.; Li, A.; Li, D.; Liu, Y.; Du, Y.; Pei, T.; Ma, T.; Zhou, C. Difference of urban development in China from the perspective of passenger transport around Spring Festival. *Appl. Geogr.* **2017**, *87*, 85–96. [[CrossRef](#)]
30. Wei, Y.; Song, W.; Xiu, C.; Zhao, Z. The rich-club phenomenon of China ' s population flow network during the country ' s spring festival. *Appl. Geogr.* **2018**, *96*, 77–85. [[CrossRef](#)]
31. Jendryke, M.; Balz, T.; Liao, M. Big location-based social media messages from China ' s Sina Weibo network: Collection, storage, visualization, and potential ways of analysis. *Trans. GIS* **2017**, *21*, 825–834. [[CrossRef](#)]
32. Peng, S.; Yang, A.; Cao, L.; Yu, S.; Xie, D. Social influence modeling using information theory in mobile social networks. *Inf. Sci. N.Y.* **2017**, *379*, 146–159. [[CrossRef](#)]
33. Sun, Q.; Guo, X.; Jiang, W.; Ding, H.; Li, T.; Xu, X. Exploring the Node Importance and Its Influencing Factors in the Railway Freight Transportation Network in China. *J. Adv. Transp.* **2019**, *2019*, 1493206. [[CrossRef](#)]
34. Ramos, A.; Lazar, M.; Filho, R.H.; Rodrigues, J.J.P.C. Model-Based Quantitative Network Security Metrics: A Survey. *IEEE Commun. Surv. Tutorials* **2017**, *19*, 2704–2734. [[CrossRef](#)]
35. Page, L.; Brin, S.; Motwani, R.; Winograd, T. *The PageRank Citation Ranking: Bringing Order to the Web*; Stanford University: Stanford, CA, USA, 1999.
36. Raghavan, U.N.; Albert, R.; Kumara, S. Near linear time algorithm to detect community structures in large-scale networks. *Phys. Rev. E* **2007**, *76*, 036106. [[CrossRef](#)] [[PubMed](#)]
37. Wu, Z.; Lin, Y.; Gregory, S.; Wan, H.; Tian, S. Balanced multi-label propagation for overlapping community detection in social networks. *J. Comput. Sci. Technol.* **2012**, *27*, 468–479. [[CrossRef](#)]
38. Newman, M.E.J. Modularity and community structure in networks. *Proc. Natl. Acad. Sci.* **2006**, *103*, 8577–8582. [[CrossRef](#)]
39. Blondel, V.D.; Guillaume, J.; Lambiotte, R.; Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech. theory Exp.* **2008**, *2008*, P1008. [[CrossRef](#)]
40. Rosvall, M.; Bergstrom, C.T. An information-theoretic framework for resolving community structure in complex networks. *Proc. Natl. Acad. Sci.* **2007**, *104*, 7327–7331. [[CrossRef](#)]
41. Rosvall, M.; Bergstrom, C.T. Maps of random walks on complex networks reveal community structure. *Proc. Natl. Acad. Sci.* **2008**, *105*, 1118–1123. [[CrossRef](#)]
42. Lancichinetti, A.; Fortunato, S. Community detection algorithms: A comparative analysis. *Phys. Rev. E* **2009**, *80*, 056117. [[CrossRef](#)]
43. Yang, Z.; Algesheimer, R.; Tessone, C.J. A comparative analysis of community detection algorithms on artificial networks. *Sci. Rep.* **2016**, *6*, 30750. [[CrossRef](#)]
44. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [[CrossRef](#)]
45. Huffman, D.A. A method for the construction of minimum-redundancy codes. *Proc. IRE* **1952**, *40*, 1098–1101. [[CrossRef](#)]
46. Guimera, R.; Amaral, L.A.N. Functional cartography of complex metabolic networks. *Nature* **2005**, *433*, 895–900. [[CrossRef](#)] [[PubMed](#)]
47. Clauset, A.; Newman, M.E.J.; Moore, C. Finding community structure in very large networks. *Phys. Rev. E* **2004**, *70*, 066111. [[CrossRef](#)] [[PubMed](#)]

48. Csardi, G.; Nepusz, T. The igraph software package for complex network research. *Int. J. Complex Syst.* **2006**, *1695*, 1–9.
49. Colizza, V.; Flammini, A.; Serrano, M.A.; Vespignani, A. Detecting rich-club ordering in complex networks. *Nat. Phys.* **2006**, *2*, 110. [[CrossRef](#)]
50. De Masi, G.; Iori, G.; Caldarelli, G. Fitness model for the Italian interbank money market. *Phys. Rev. E Stat. Nonlinear Soft Matter. Phys.* **2006**, *74*, 066112. [[CrossRef](#)] [[PubMed](#)]
51. Zhou, S.; Mondragón, R.J. The Rich-Club Phenomenon in the Internet Topology. *IEEE Commun. Lett.* **2004**, *8*, 180–182. [[CrossRef](#)]
52. Opsahl, T.; Colizza, V.; Panzarasa, P.; Ramasco, J.J. Prominence and Control: The Weighted Rich-Club Effect. *Phys. Rev. Lett.* **2008**, *101*, 168702. [[CrossRef](#)]
53. Opsahl, T. *Structure and Evolution of Weighted Networks*; Queen Mary College, University of London: London, UK, 2009.
54. Long, Y.; Wu, K. Shrinking cities in a rapidly urbanizing China. *Environ. Plan. A* **2016**, *48*, 220–222. [[CrossRef](#)]
55. Guo, D. China's Tibetan population and population in Tibet. *China Popul. Today* **1996**, *13*, 7–8. [[PubMed](#)]
56. Miao, Y. Analysis of population change in Hulunbeier City. *Inn. Mong. Stat.* **2016**, 59–61. [[CrossRef](#)]
57. Liu, H.; Du, G. Spatial-temporal pattern of China's economic development and its dynamic evolution: Based on city level DMSP/OLS night-time lights data. *Chin. J. Popul. Sci.* **2017**, *37*, 17–29.
58. Liu, H.; He, L. Regional disparity in China and its evolution (2000–2012)—Re-examination based on DLI. *Rev. Econ. Manag.* **2015**, 141–146. [[CrossRef](#)]
59. Guan, X.; Fang, C.; Luo, K. Regional economic development disparity of China: An application of spatial field. *Sci. Geogr. Sin.* **2012**, *32*, 1055–1065. [[CrossRef](#)]
60. Chang, L. The Current Status and the Development of Tibetan Culture Industry under the Global Horizon. *J. Tibet Univ.* **2009**, *24*, 32–37. [[CrossRef](#)]
61. Wu, Y.S. Taiwan's domestic politics and cross-strait relations. *China J.* **2005**, 35–60. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).